

SBORNÍK

prací účastníků vědecké konference doktorského studia

Fakulta informatiky a statistiky

Vysoké školy ekonomické v Praze



**Vědecká konference se uskutečnila dne 13. února 2014
pod záštitou děkana FIS doc. RNDr. Luboše Marka, CSc.**

Sestavení sborníku

prof. Ing. Petr Doucek, CSc.

proděkan pro vědu a výzkum

© Vysoká škola ekonomická v Praze
Nakladatelství Oeconomica – Praha 2014
ISBN 978-80-245-2010-0

OBSAH

Předmluva	5
Využití modelů Systémové dynamiky při řešení vybraných problémů manažerské komunikace.....	7
<i>Martin Dalihod</i>	
Studium vlivu behaviorálních a sociálních faktorů v organizačních procesech řízení kvality softwaru a dodávek informačních systémů.....	16
<i>Michal Doležel</i>	
Řízení Informační bezpečnosti na základě standardního chování.....	28
<i>Martin Dvořák</i>	
Přínosy agilních metodik řízení ICT projektů.....	35
<i>Jan Juříček</i>	
Využití metodiky COBIT 4.1 při penetračním testování bezpečnosti IS	43
<i>Tomáš Klíma</i>	
Alokace lidských zdrojů v projektovém řízení	50
<i>Roman Nedzelský</i>	
Agile Outsourcing of Software Development.....	57
<i>Aziz Ahmad Rais</i>	
Přispívá internet k rozvoji tržního systému?	66
<i>Václav Řezníček</i>	
Přehodnocení přístupů komerčních subjektů k řízení marketingových aktivit v prostředí služeb na internetu.....	73
<i>Zdeněk Smutný</i>	
Metriky, monitoring a řídicí proces Data Governance	83
<i>Martin Vacek</i>	
Příprava modelu multimediální komunikace	93
<i>Zdeněk Vondra</i>	
Vývoj nástroje pro diskrétní simulaci	104
<i>Karel Charvát</i>	
Fuzzy approach for portfolio selection problem.....	114
<i>Maria Kobzareva</i>	
Analýza vztahů exportu Ruska, ceny ropy a dalších veličin	124
<i>Elena Kuchina</i>	
GARCH(1,1), metoda maximální věrohodnosti a počet pozorování	134
<i>Ing. Peter Princ</i>	
Nekooperativní vícekritériální hry	144
<i>Michaela Tichá</i>	
Demografický a ekonomický pohled na populační stárnutí.....	151
<i>Kornélia Cséfalvaiová</i>	
Jaký vliv budou mít modely používané pro extrapolaci křivky úmrtnosti na hodnotu střední a normální délky života?	161
<i>Petra Dotlačilová</i>	

Nadúmrtnost mužů v ČR.....	167
<i>Jana Langhamrová</i>	
Metódy neparametrickej interpolácie dát.....	177
<i>Tomáš Marcinko</i>	
Statistické analýzy regionálních diferencí v oblasti udržitelného rozvoje a kvality života	187
<i>Ludmila Petkovová</i>	
Pracovní migrace v České republice.....	194
<i>Martina Šimková</i>	
Modelování demografických časových řad pro potřeby demografických projekcí	204
<i>Ondřej Šimpach</i>	
Porovnání nových přístupů ve shlukování nominálních dat	214
<i>Zdeněk Šulc</i>	

Předmluva

V letošním roce se konal pod gescí děkana Fakulty informatiky a statistiky devatenáctý ročník semináře „Den doktorandů“. Tento seminář se stal tradiční platformou, kde doktorandi všech oborů doktorského studia fakulty prezentují výsledky své vědecké práce. Pro mnohé z nich je to první vystoupení před odbornou veřejností, na němž získávají cenné zkušenosti v oblasti formulace svých názorů a hypotéz, jejich presentace a argumentace na jejich obhajobu pro svoji další vědeckou práci ve studovaných oborech. I v letošním roce byly příspěvky rozděleny do tří sekcí podle studovaných oborů doktorského studia na Fakultě informatiky a statistiky. Na devatenáctý „Den doktorandů“ se celkem přihlásilo 24 účastníků se svými příspěvky. Z toho na oboru „Aplikovaná informatika“ to bylo 11 příspěvků, na oboru „Ekonometrie a operační výzkum“ 5 příspěvků a na oboru „Statistika“ 8 příspěvků. Semináře se zúčastnili studenti obou forem studia, jak presenční, tak i kombinované.

Nedílnou součástí „Dne doktorandů“ je i práce hodnotících komisí, které pečlivě sledují jednotlivá vystoupení a potom po oborech vybírají nejlepší práce k ocenění. Hlavními kritérii pro jejich rozhodování byly zejména náročnost a aktuálnost zpracovaného tématu, přístup k řešení vybraného problému, způsob použití metodiky, úroveň práce s reálnými daty a v neposlední řadě i schopnost prezentovat a argumentačně obhájit výsledky své vědecké práce v diskusi jak se svými vrstevníky, tak i se staršími kolegy. Ti nejlepší z nich získávají prestižní „Cenu děkana FIS“, s níž je spojena i symbolická finanční odměna.

Za práci v komisích chci poděkovat všem jejím členům, kteří pracovali pod vedením předsedů - doc. Ing. Vojtěcha Svátka, Dr. (obor Informatika), prof. Ing. Romana Huška, CSc. (obor Ekonometrie a operační výzkum) a prof. Ing. Hany Řezankové, CSc. (obor Statistika). Že se jednalo o nelehkou práci, ukazuje i fakt, že ve všech vyhlášených kategoriích byla udělena dělená přední místa. Komise se své nelehké úlohy zhostily na výbornou.

V letošním roce získali ceny za nejlepší příspěvky v jednotlivých kategoriích následující studentky a studenti:

Studijní program – Aplikovaná informatika

obor Aplikovaná Informatika

1. místo: Neuděleno

2. místo: **Ing. et Ing. Michal Doležel:** Studium vlivu behaviorálních a sociálních faktorů v organizačních procesech řízení kvality softwaru a dodávek informačních systémů

Ing. Tomáš Klíma: Využití metodiky COBIT 4.1 při penetračním testování bezpečnosti IS

3. místo: **Mgr. Ing. Zdeněk Smutný:** Přehodnocení přístupů komerčních subjektů k řízení marketingových aktivit v prostředí služeb na internetu

Ing. Zdeněk Vondra: Příprava modelu multimediální komunikace

Studijní program – Kvantitativní metody v ekonomice

obor Ekonometrie a operační výzkum

1. místo: **Ing. Elena Kuchina:** Analýza vztahů exportu Ruska, ceny ropy a dalších veličin

Ing. Peter Princ: GARCH(1,1), metoda maximální vierohodnosti a počet pozorování

2. místo: Neuděleno

3. místo: **RNDr. Ing. Michaela Tichá:** Nekooperativní vícekritériální hry

obor Statistika

1. místo: **Ing. Zdeněk Šulc:** Porovnání nových přístupů ve shlukování nominálních dat

2. místo: **Mgr. Ing. Tomáš Marcinko:** Metódy neparametrickej interpolácie dát

Ing. Ondřej Šimpach: Modelování demografických časových řad pro potřeby demografických projekcí

3. místo: **Ing. Martina Šimková:** Pracovní migrace v České republice

Čestné uznání za příspěvek z oblasti ekonomické a sociální statistiky obdržela Ing. Ludmila Petkovová.

Oceněným studentům doktorského studia upřímně blahopřeji a doufám, že získané zkušenosti uplatní při své další práci, ať už vědecké nebo v praxi. Uznání také patří všem vědeckým a pedagogickým pracovníkům FIS – školitelům doktorandů, kteří se Dne doktorandů zúčastnili a svým vedením a radami byli nápomocni při zpracování příspěvků. Svojí účastí a dotazy v diskusi pak vytvořili na semináři kreativní atmosféru celého setkání.

Zvláštní poděkování pak patří studijní referentce doktorského studia paní Jitce Krajičkové, díky níž byl seminář skvěle organizačně zajištěn, a Mgr. Lee Nedomové za obětavou práci při editaci a sestavení tohoto sborníku.

Prof. Ing. Petr Doucek, CSc,
Proděkan pro vědu a výzkum

Využití modelů Systémové dynamiky při řešení vybraných problémů manažerské komunikace

Martin Dalihod

xdalm03@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: doc. Ing. Stanislava Mildeová, CSc., (mildeova@vse.cz)

Abstrakt: Článek se zabývá aplikací modelů Systémové dynamiky do manažerské praxe. V současném světě, kdy vedení hospodářských organizací již vnímá stále rostoucí význam systémového myšlení pro řešení problémů spojených s narůstající komplexitou informačních systémů dané organizace, prozatím nedošlo ke změně lineárního paradigmatu, které je pevně ukotveno v idiosynkratickém modelu jednotlivých manažerů. Cílem článku je upozornit na možnost sdílení mentálního modelu směřující k nápravě idiosynkratického modelu, který je součástí aspektů manažerské komunikace a který je zohledňován při manažerském rozhodování.

Pro účely článku byly vybrány tři problémy tak, aby pokryly významné oblasti manažerské komunikace. Konkrétně se jedná o oblast Vize, Projektového řízení a Auditů.

Ke svému rozhodování člověk v dnešní době velmi často využívá tzv. systémy na podporu rozhodování (Decision Support Systems - DSS), kdy dochází ke konfrontaci mentálních modelů formou dílčích řešení prostřednictvím informačních technologií. Informační technologie umožňují kladení dotazů formou experimentů s reálnou podobou mentálních modelů. V praxi se tyto modely využívají jako speciální, nebo univerzální (obecné) systémy, které pomáhají jejich uživatelům (manažerům) při realizaci řídicích a rozhodovacích činností. Tyto modelové orientované systémy nevyužívají ke své potřebě celý rozsah objemných podnikových databází. Jejich podstatou je samotná tvorba modelu, simulace, a smyslem je pak pomoci analyzovat pákové efekty či vazby mezi proměnnými (ty se vyberou jako vzorek z databáze – a to nejprve pro účely verifikace modelu a následně pro tvorbu predikcí). Modely Systémové dynamiky, jejichž cílem je řešit vybrané problémy manažerské komunikace, jsou v článku zpracovány právě do podoby návrhů systémů pro podporu rozhodování.

Článek by měl přispět k odborné multidisciplinární diskusi a poskytnout některé podněty pro úvahy k otázce: Jakou konkurenční výhodou mohou přinést v článku popsané aplikace DSS do současného prostředí hospodářské organizace, pokud budou použity v manažerské komunikaci?

Klíčová slova: Systémová dynamika, Systémy pro podporu rozhodování, Mentální model, Systémové myšlení, Systém, manažerská komunikace, konkurenční výhoda.

Úvod

V současné době se v manažerské komunikaci setkáváme se situacemi, kde tradiční přístupy již nepři-
náší požadované výsledky. Autor článku vnímá jako významný nedostatek při řešení současných problémů manažerské komunikace nedostatek systémového přístupu, omezené zaměření na kontext a také neschopnost správně pochopit jiné systémy v hospodářské organizaci, zejména měkké systémy. Systémové myšlení se velmi zřídka praktikuje v oblasti manažerské komunikace, a to přesto, že je naprosto nezbytné pro systémové řešení problémů.

Cílem článku je přispět k posílení systémového myšlení v předmětné oblasti manažerské komunikace prostřednictvím modelů Systémové dynamiky. Jádrem Systémové dynamiky, vytváření explicitních modelů, které jsou externím rozšířením našich vlastních mentálních modelů. Modely Systémové dynamiky jsou v článku zpracovány do podoby modelově orientovaných návrhů systémů pro podporu rozhodování, jejichž cílem je řešit vybrané problémy manažerské komunikace. Pro účely článku byly vybrány tři problémy pokrývající významné oblasti manažerské komunikace. Konkrétně se jedná o oblast Vize, Projektového řízení a Auditů. Pro tyto zkoumané problémy jsou vypracovány případové studie.

Pro tvorbu modelů Systémové dynamiky a návrh DSS byl použit software Insight Maker.

Problémy manažerské komunikace

V průběhu celého života člověk skrze interakce se svým okolím neustále validuje svůj mentální model. Prostřednictvím komunikace dochází ke kooperaci i k soupeření, k vzájemnému vnímání. Explicitní vyjádření mentálního modelu jako systému vede k percepci, která vede nejen k pochopení asociací daného systému, ale i k efektivnější výměně sdělení v rámci sociálního styku [16].

Velké množství ztrát podniků není způsobeno tím, že by byl nedostatek signálů včasného varování nebo že jsou agilnější, ale z důvodů myšlení manažerů. Problémem je ignorování signálů a vlastní přesvědčení manažerů. Významné studie poukazují na fakt, že pokud je realita v rozporu s přesvědčením, tak nakonec při rozhodování rozhodne přesvědčení. Manažeri často z důvodů naprostého neporozumění podstatě problému řeší riziko jako funkční problém a neberou v potaz, že riziko prochází průřezově všechny podnikové funkce a obvykle se netýká jen té jedné, která je často jen viditelným vnějším symptomem (signálem) mnohem významnějšího rizika, jehož dopad se projeví v delším horizontu. Z výzkumů provedených v ČR byla zjištěna absence včasného varování v informačních systémech. V současném prostředí ČR (turbulentní podnikatelské prostředí) má včasné rozpoznání signálů velký význam v konkurenčním boji [14].

Pro potřeby článku byly vybrány tyto problémy:

- Komunikace závěrů auditu IT procesu
- Komunikace generálního ředitele pro sdílení vize hospodářské organizace
- Komunikace stavu projektu rozvoje informačního systému

Problémy byly zvoleny tak, aby pokryly významné oblasti manažerské komunikace, a to z hlediska jak je vnímá autor článku.

Systémové myšlení a modely Systémové dynamiky

Pojem "systém" je stěžejním pojmem obecné teorie systému a je jedním z nejrozšířenějších pojmů vědy. Mnoho vědních disciplín používá tento pojem různým způsobem pro řešení nejrůznějších problémů. [2] Ty rozdíly různých řešení jsou dány především tradicí a cíly jednotlivých věd. V článku práci je pojem systém chápán jako dynamický systém, který se v čase vyvíjí a je nelineární. V lineárních systémech malá změna způsobuje pouze malou změnu v kvalitativním chování systému. Naproti tomu v nelineárním systému může malá změna způsobit dramatickou změnu kvalitativního chování systému. Účelem systémového modelu je pak odhalovat stávající problémy daného systému a prostřednictvím pákového efektu je možné systém stimulovat pomocí preventivních, detektivních a následných opatření. Model pak představuje efektivní přístup k zvýraznění charakteristických aspektů systémů, tj. jejich strukturu, vnitřní systémovou interakci mezi složkami, kauzální vztahy, vstupy a výstupy, tok materiálu a dat uvnitř systému a logiku rozhodovacích pravidel [1, 7, 17].

Aktuální dění, které je možné sledovat prostřednictvím sdělovacích prostředků, nabádá k zamyšlení, co vede management hospodářské organizace k tomu, že umožní, aby současná krize napáchala takové škody. Proč se neustále opakuje situace, kdy si hospodářská organizace prostřednictvím svého vedení nepřipouští, že by pro ni zdánlivě nepodstatný problém mohl přerůst v krizi [5]. Doktorand vnímá současný vývoj ve zkoumané oblasti tak, že stále přetrvává trend, kdy se problémy stále řeší ve chvíli, kdy se objeví, a to jako jednotlivé události. Vedení hospodářských organizací se stále nenaučilo prostřednictvím Systémové dynamiky vnímat nejen jednotlivé situace, ale i vzory chování v čase. V manažerské praxi stále chybí aplikace Systémové dynamiky k modelování problémů a tím pochopení struktur systémů. Vedení hospodářské organizace si neuvědomuje, že veškeré dynamické chování je důsledkem struktury systému a že prostřednictvím modelování je možné efektivním způsobem komunikovat preventivní a nápravná opatření.

Současný vývoj manažerské komunikace prozatím nezná jeden z významných pojmů kognitivní vědy, který propojuje systémovou dynamiku se systémovým myšlením, a to tzv. mentální model, a to i přesto, že brožury manažerské komunikace již s tímto pojmem nevědomky operují prostřednictvím tzv. neuro-lingvistického programování, jež se zabývá různými způsoby vnímání a jejich vlivem na úspěšnou komunikaci. Tvůrci teorie neurolingvistického programování rozlišují dvě jazykové struktury. Hlubokou strukturu jazyka a povrchovou strukturu jazyka. Významný problém, který je v současné době stále

přehlížen představuje fakt, že manažeři pro každodenní komunikaci používají povrchovou strukturu jazyka, kdy z hluboké struktury jazyka jen vyberou některé informace a značnou část vypouštějí. Jednoduše zobecňují. Mentální informace existují na velmi hluboké úrovni v naší nervové soustavě. Pro zmapování mentálních modelů se dnes využívají tzv. příčinné smyčkové diagramy, které se dále prostřednictvím softwarových nástrojů převádí na tzv. diagramy stavů a toků a tak uvádějí struktury do souvislosti s jejich chováním. Chování struktury se projeví v průběhu simulace modelu [18, 6, 10, 8].

Mezi nejvýznamnější softwarové nástroje, které se využívají k tvorbě modelů Systémové dynamiky, v současné době patří Vensim a Insight maker. V době, kdy byla Systémová dynamika vyvinuta, byl využíván program DYNAMO, který byl však nepříliš intuitivní a podobal se spíše programovacímu jazyku. Převratem byl software Powersim, který je intuitivní a snadno použitelný. Současný software Vensim je možné označit jako vylepšený software Powersim, který již byl v průběhu let plně odlazen pro potřeby tvorby modelů Systémové dynamiky a pro potřeby současné manažerské praxe. Současná doba však zažívá velmi významnou změnu, kterou představuje software Insight Maker, který umožňuje oproti původnímu trendu, kdy tvorba modelů byla pouze v kompetenci odborníků na Systémovou dynamiku, možnost zapojení do tvorby modelů Systémové dynamiky širokou veřejnost. Insight Maker, jehož autorem je významný systémový myslitel současné doby Gene Bellinger, již není závislý na platformě a je k dispozici zcela zdarma. Prostřednictvím přidání obrázků do příčinných smyčkových diagramů, umožnění sdílení vytvořených modelů a uveřejnění četných tutoriálů, které umožňují široké veřejnosti nahlédnout pod pokličku, kam viděli doposud jen odborníci na Systémovou dynamiku, představuje Insight maker ještě více intuitivní a snadno použitelný nástroj, který posouvá Systémovou dynamiku do zcela nové etapy a dává příležitost i pro ty, kteří nejsou počítačovými nebo matematickými experty, aby komunikovali své mentální modely.

Grafické znázornění je prováděno pomocí nástrojů Systémové dynamiky, která poskytuje perspektivní sadu nástrojů, které nám umožňují pochopit strukturu a dynamiku komplexních systémů. Tyto nástroje nám umožňují vytvářet simulace, kde prostor a čas může být komprimován a čas zpomalen a my tak můžeme vidět dlouhodobý vedlejší efekt našeho rozhodnutí, což vede ke zrychlení našeho učení a my tak dosáhneme rozvoje našeho chápání komplexních systémů v návaznosti na jejich strukturu [13].

Současná metodika tvorby modelů Systémové dynamiky

Současná metodika využívá dva nástroje příčinný smyčkový diagram a diagram stavů a toků.

Příčinný smyčkový diagram (CLD causal loop diagram)

Zpětná vazba je v článku znázorněna pomocí příčinného smyčkového diagramu, prostřednictvím něhož je možné zachytit oba typy zpětné vazby. Tvorba příčinného smyčkového diagramu představuje proces, který začíná identifikací vztahu mezi jednotlivými dvojicemi proměnných. Pokud změna jedné proměnné produkuje změnu ve stejném směru, v druhé proměnné je vztah pozitivní. Pokud změna v druhé proměnné běží v opačném směru, je vztah definován jako negativní. Proměnné jsou spojeny dohromady, aby vytvořily smyčky odezvy systému [4]. Článek poukazuje na fakt, že příčinný smyčkový diagram zjednodušuje transformaci slovního popisu do zpětnovazební struktury [11].

Diagram stavů a toků (SFD stock and flow diagram)

SFD představuje převedení statických příčinných diagramů, které zachycují dynamiku systému do podoby dynamických modelů a simulátorů [15].

Posláním modelů (vybraných problémů článku) je pochopit systém jako celek prostřednictvím nalezení jeho tzv. „leverage points“ (dále jen LP), jejichž změnou je možné systém ovlivňovat [9].

Modely jsou vytvářeny podle zásad etapizace zásad tvorby modelů Systémové dynamiky dle [3, 15]:

- 1) **Definice účelu** – umožňuje odlišení podstatného od nepodstatného, určit hranice a časový horizont systému, pochopit problém; stanovujeme historické chování proměnných => referenční módy; modelujeme problém a ne celý systém.
- 2) **Formulace dynamických hypotéz** – předpoklad pro učení se; zaměřujeme se na chování systému jako celku pomocí jeho endogenních charakteristik - hledáme příčinnou zpětnovazební

strukturu (vzor chování). Nástroje: diagramy hranic systému, diagramy subsystémů, příčinné smyčkové diagramy, diagramy stavů a toků.

- 3) **Formulace simulačního modelu** – základ pro odvození důsledků; kvantifikace vlastností (SW nástroje umožňují vytvářet modely na základě grafických diagramů stavů a toků); formulace matematických vztahů mezi prvky modelu; testování konzistence
- 4) **Testování** – chování modelu ve vztahu k referenčním módům, robustnosti modelu při extrémních podmínkách; citlivost na počáteční podmínky; míra neurčitosti;
- 5) **Návrh a posouzení politik** – návrh nových reálně použitelných strategií a struktur vedoucích k nápravě problému; posouzení možných nastavení; If-then analýzy politik a jejich nechtěných efektů; posouzení vzájemných interakcí mezi různými politikami (protichůdné a navzájem se kompenzující vlivy, posílení).

Iterativnost procesu tvorby modelů – pátým krokem proces nekončí, vše se neustále opakuje; často je třeba se vracet k předchozím krokům.

Současným trendem většiny firem v návaznosti na ekonomickou krizi se stal oblíbeným výrazem „cost cutting“. Firmy si velmi dobře uvědomují, že pokud budou mít ve svých řadách odborníky na systémové myšlení, tak ušetří výdaje za často nesmírně drahé analýzy konzultantských firem, které končí jako archy založené v šuplicích manažerů bez dalšího využití. Nevýhodou však představuje fakt, že schopnost systémového myšlení vyžaduje nejprve skutečné pochopení systémového myšlení na straně organizace a následné zapojení do praktického užívání. K tomu by mohlo přispět to, že model lze vytvářet formou team buildingu, což je nový světový trend v oblasti tvorby modelů Systémové dynamiky. Účastníky team buildingu by pak měl být výběr z řad top manažerů i řadových zaměstnanců. Nejen že takto dojde k pochopení fungování systému, ale i k odbourání často zbytečné byrokracie.

Návrh DSS

Manažerskou komunikaci je z pohledu vybraných problémů možné rozdělit na tři stěžejní oblasti z pohledu času (viz obrázek 1).



Obrázek 1

První problematickou oblastí manažerské komunikace je definice vize, která představuje rámec pro řízení hospodářské organizace. V návaznosti na danou vizi je třeba nastavit veškeré komunikační kanály tak, aby byly v souladu se stanovenou vizí. V této oblasti představuje model Systémové dynamiky ideální nástroj pro komunikaci návrhu vize napříč hospodářskou organizací. Korektní nastavení datových toků tak, aby signály včasného varování byly identifikovány včas a v reálném čase, je první z významných úkolů každého manažera. V současné době není tento způsob v praxi manažerské praxe aplikován. Vize není centralizována a není komunikována prostřednictvím modelu Systémové dynamiky. Problém představuje fakt, že vize je komunikována prostřednictvím nesystémových a často decentralizovaných pokynů, případně metodikou lineárního charakteru. Autor článku považuje model Systémové dynamiky jako vhodný komunikační nástroj, který zabrání zatajování či účelovému generování nereálných dat, které jsou zapotřebí k verifikaci stavu o naplňování zvolené vize.

Druhou problematickou oblast manažerské komunikace představuje samotný životní cyklus projektu. Autor článku zde vnímá entitu projektu jako proces, který má zmapovány všechny klíčové milníky a pro každý milník jsou zcela jasně stanovena pravidla ohledně aktivit a dokumentace. Vnímané problémy jsou způsobeny opět nesystémovým přístupem, které mají dlouhodobé negativní dopady. Autor článku

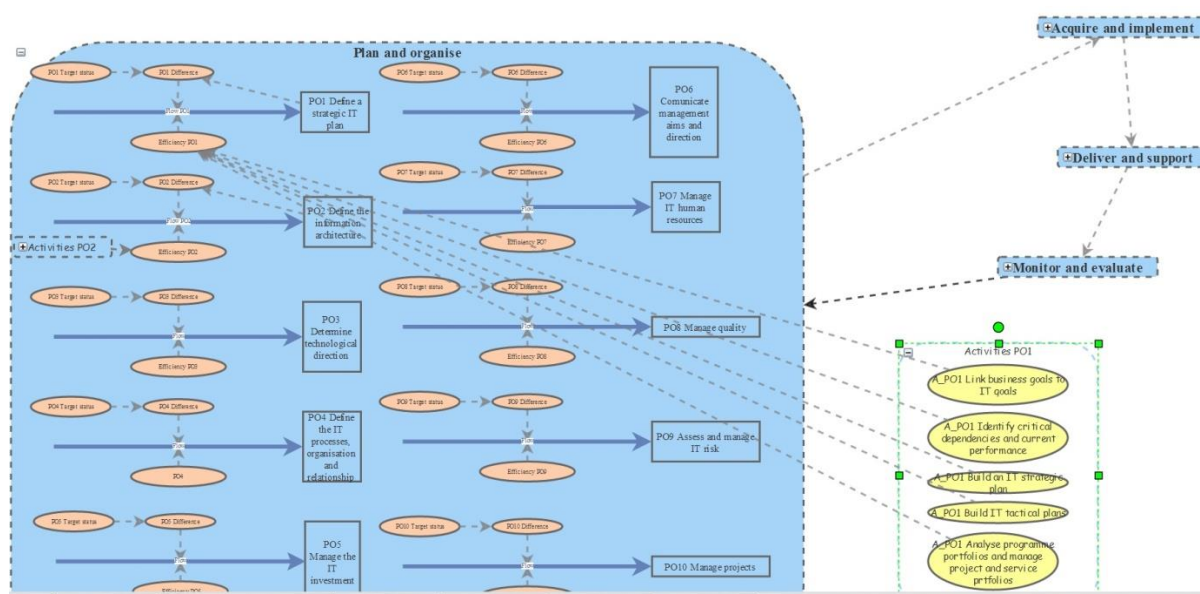
se domnívá, že současné problémy projektového řízení je možné odstranit při aplikaci modelu Systémové dynamiky.

Poslední problematickou oblastí je oblast auditu, tedy kontroly. Hospodářskou organizaci je třeba vnímat jako systém, který je možné kontrolovat pomocí Systémového myšlení nejprve z jeho statického pohledu, tedy z pohledu jeho vize a následně z pohledu dynamického, tedy z pohledu kvality životního cyklu a dokumentace projektů. V současné době se audit zaměřil spíše na datovou bezpečnost, nežli na optimalizaci procesů hospodářské organizace, což je velmi závažný problém.

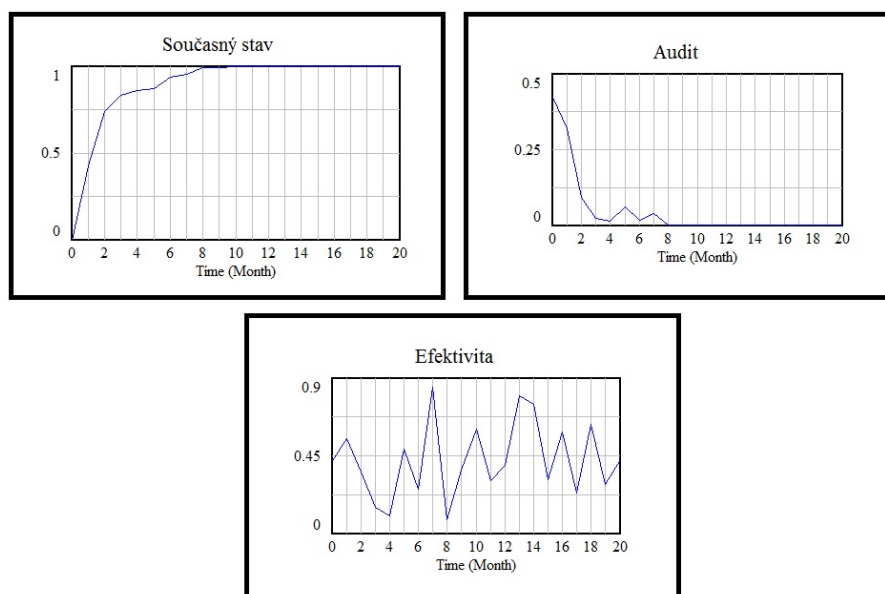
V dalších kapitolách článek rozebere jednotlivé DSS z pohledu univerzálního (obecného) systému, které pomáhají jejich uživatelům (manažerům) při realizaci řídicích a rozhodovacích činností. Speciální využití DSS a jejich průběžné ladění v průběhu jednotlivých iterací není součástí tohoto článku.

Návrh DSS pro komunikaci závěrů auditu IT procesu

Model číslo jedna pomáhá naplňovat jedno z nejdůležitějších pravidel managementu, a to "Kdo řídí, ten kontroluje". Pomocí aplikace systémového myšlení je demonstrována alternativa komunikace závěrů interní kontroly oproti nesystémovému lineárnímu přístupu.



Obrázek 2

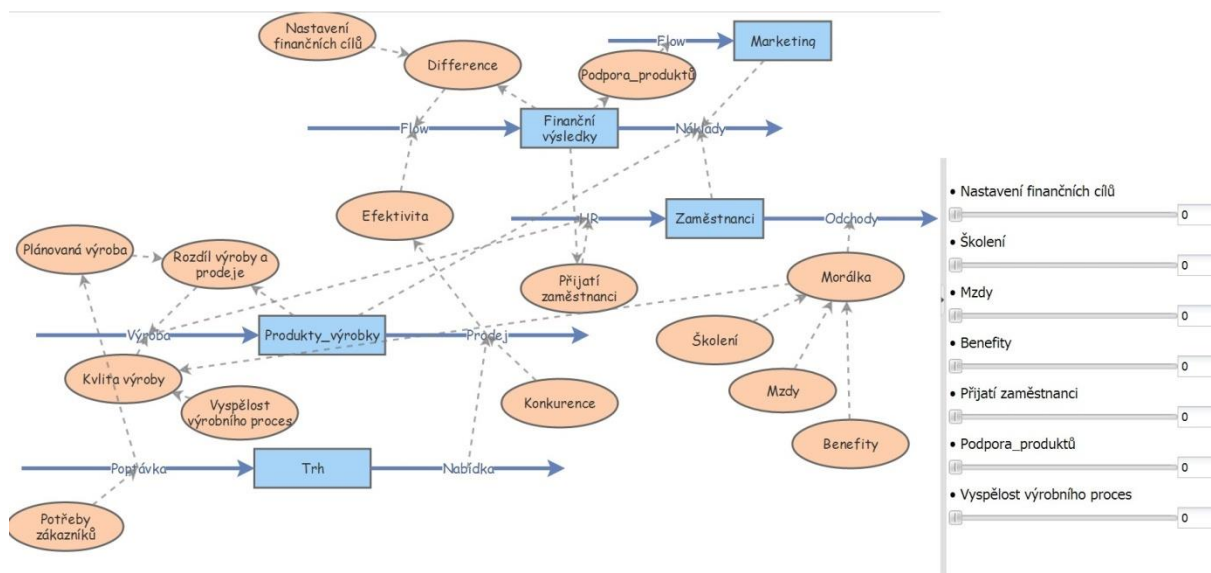


Obrázek 3

(Obrázek 2 a 3) DSS komunikace závěrů auditu IT procesu z pohledu univerzálního (obecného) systému Návrh DSS komunikace závěrů auditu IT procesu se snaží plně využít doporučení metodiky COBIT (sada osvědčených postupů) k řízení procesů a postupů v oblasti infrastruktury informačních a komunikačních technologií, a to s měřitelnými kontrolami. Toto DSS nemá, oproti následujícím příkladům, možnost interakce s modelem prostřednictvím ovlivňování vybraných proměnných. DSS představuje v tomto případě "engine" DSS (jádro DSS, které je tvořeno převedením příčinného smyčkového diagramu na diagram stavů a toků), který převzal COBIT (sada osvědčených postupů) a provázal je online s prováděnými kontrolami interního auditu (obrázek 2). Uživatelé (manažeři) mají pak možnost prostřednictvím rozhraní DSS (obrázek 3) sledovat efektivitu aplikace COBIT (sada osvědčených postupů) do praxe. DSS jim však poskytuje nejen možnost sledování iterace jednotlivých IT procesů, ale zejména sledovat iterace v procesech, které jsou na sobě závislé dle metodiky COBIT[12].

Návrh DSS pro komunikaci generálního ředitele pro sdílení vize hospodářské organizace

Předtím, než je možné chod organizace kontrolovat, je zapotřebí nastavit řídicí mechanismus, který bude validovat nejdůležitější proměnné tak, jak je vnímá generální ředitel hospodářské organizace. Pro tyto účely byl vytvořen model Systémové dynamiky, který je navržen tak, aby mohl podpořit komunikaci generálního ředitele hospodářské organizace. Výsledný model slouží v počátcích jako podklad pro navržení cílového konceptu pro komunikaci potřeby sběru dat důležitých pro potřeby přijímáním rozhodnutí nejvyššího vedení společnosti. Následně model slouží jako manažerský simulátor pro validaci klíčových rozhodnutí.



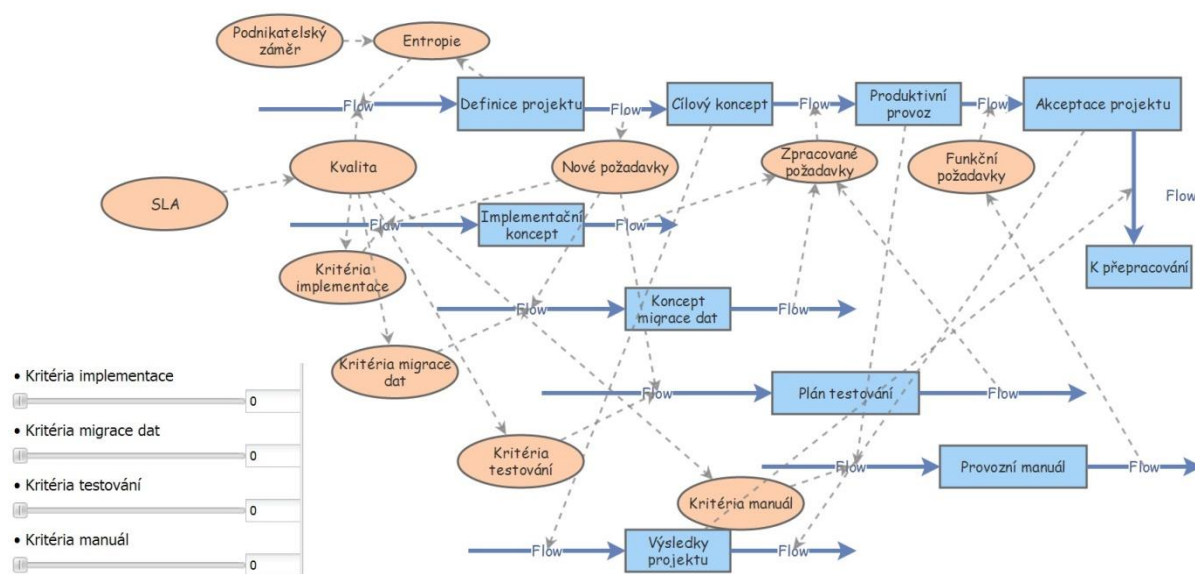
Obrázek 4

(Obrázek 4) DSS pro komunikaci generálního ředitele pro sdílení vize hospodářské organizace z pohledu univerzálního (obecného) systému Návrh DSS v tomto případě již umožňuje interakci s modelem. DSS umožňuje v návaznosti na nastavení finančního očekávání, které "engine" DSS zohlední v porovnání s aktuálními parametry hospodářské organizace a tržními výsledky v podobě finančních výsledků za danou iteraci. Uživatel (generální ředitel) má pak možnost simulovat svá rozhodnutí prostřednictvím interakce s proměnnými. Tedy jaký dopad bude mít investice do školení zaměstnanců, mezd, příjmu nových zaměstnanců, marketingové podpory a modernizace výroby. Výsledné simulace, interakce s DSS v návaznosti na skutečnost a její odladění s modelem pak povedou k efektivnější komunikaci strategických rozhodnutí a nastavení interních politik (například politice zaměstnanosti).

Návrh DSS pro komunikaci stavu projektu rozvoje informačního systému

Třetí model má za cíl monitorovat kvalitu realizace IT projektů dle předem definovaných kritérií. V úplných počátcích životního cyklu projektu, tedy v průběhu tvorby definice projektu, slouží model jako podklad pro analýzu, zda definice projektu je v souladu se strategií korporace a zda pokrývá všechny

procesy určené zadavatelem. Z cílového konceptu, který autor považuje za klíčový dokument celého projektu, jsou pro potřeby modelu nejprve extrahovány definované vlastnosti a akceptační kritéria projektu. Tvorba modelu prověří, zda byla zpracována procesní analýza podle standardní metodiky, zvolené dodavatelem, např. metodika UML. Z cílového konceptu jsou dále extrahovány všechny požadavky zadavatele s důrazem na detailní popis procesu a podrobnost technického řešení. V případě nejasností jsou bezodkladně komunikovány eskalace na manažerské rozhodnutí zadavatele. Pro akceptaci modelu je nastaveno kritérium maximálně 2 % nejasných „bílých míst“, která se ujasní během implementace, například legislativní změny, které prozatím nejsou známy etc. Rozsah projektu musí být v souladu s TCO a cílový koncept musí být v souladu s harmonogramem. V dalších fázích životního cyklu projektu model pomáhá sledovat odchylky harmonogramu od plánu v detailu definice dílčích cílů (milestone). Dále umožňuje sledovat finanční „očekávky“ projektu a kontrolovat finanční výstupy. Speciální/mimořádná pozornost je věnována procesu testování, protože autor článku vnímá tuto dílčí část za další velmi důležitý milník v projektu, který je často podceňován. Ve stanovené frekvenci model vyhodnocuje realizaci testů z hlediska časového a kvalitativního. Ověřuje, zda jsou testy prováděny dle harmonogramu, obsahující výsledky testů. Model umožňuje v průběhu fáze testování zejména analyzovat návrhy nedostatků dílčích opatření.



Obrázek 5

(Obrázek 5) DSS komunikaci stavu projektu rozvoje informačního systému Návrh DSS v tomto případě opět umožňuje interakci s modelem. DSS umožňuje nastavit kritéria, podle kterých se v daném projektu bude porovnávat řídicí dokumentace s realitou. Proměnné v podobě kritérií mají v návaznosti na řídicí dokumentaci zafixovány alternativy kritérií (např. implementační koncept je v souladu s harmonogramem projektu - odchylka max 5; plán testování obsahuje plán pro testování všech položek cílového konceptu - odchylka 0%, etc.), u kterých je možné prostřednictvím DSS v jednotlivých iteracích sledovat jejich plnění a odchylky. Při speciálním využití modelu (viz kapitola Návrh DSS) budou vyhodnocovány odchylky v celém životním cyklu projektu rozvoje informačního systému (viz model obrázek 5) a přijímána nápravná opatření. DSS v tomto případě nejen že umožní komunikovat stav projektu, ale pomůže odhalit problematická místa životního cyklu a nastavit kritéria tak, aby výsledný produkt projektu (informační systém) byl co možná nejvíce v souladu s podnikatelským záměrem a byly minimalizovány požadavky na přepracování.

Závěr

Článek podává důkaz o tom, že modely Systémové dynamiky, pokud jsou rozšířeny do podoby DSS a jejich "engine" DSS je vytvořen podle osvědčených postupů, ať už se jedná o metodiku COBIT a nebo empirické poznatky v podobě řídicí dokumentace, je možné využít jako systémový nástroj, který vede k optimalizaci manažerské komunikace. Závěrem článku je tedy to, že Systémová dynamika s využitím

informačních technologií umožňuje modelovat chování reálného systému v čase a je tedy významným nástrojem, který je možné využít při řešení problémů manažerské komunikace.

Článek by měl přispět k odborné multidisciplinární diskusi a poskytnout některé podněty pro úvahy k otázce: Jakou konkurenční výhodu mohou přinést v článku popsané aplikace DSS do současného prostředí hospodářské organizace, pokud budou použity v manažerské komunikaci?

Článek by měl dále inspirovat k dalším možnostem, jak rozšířit v článku zmíněné DSS a zejména povzbudit úvahy o jejich speciálním využití.

Literatura

- [1] DAELLENBACH, Hans G, D MCNICKLE a Shane DYE. Management science: decision making through systems thinking. 2nd edition. ISBN 978-023-0316-478.
- [2] EXNAROVÁ, Anna, DALIHOD, Martin, MILDEOVÁ, Stanislava. Measuring systems thinking ability. In: Efficiency and Responsibility in Education. Praha, 09.06.2011 – 10.06.2011. Prague : Czech University of Life Sciences in Prague, 2011, s. 66–74. ISBN 978-80-213-2183-0.
- [3] FORRESTER, Jay W. Principles of systems. Waltham, MA: Pegasus Communications, Inc. ISBN 18-838-2341-2.
- [4] GOODMAN, Michael R. Study notes in system dynamics. Cambridge, Mass.: Wright-Allen Press, ISBN 09-147-0000-6.
- [5] CHALUPA, Radek. Efektivní krizová komunikace: pro všechny manažery a PR specialisty. 1. vyd. Praha: Grada, 2012, ISBN 978-80-247-4234-2.
- [6] JIŘINCOVÁ, Božena. Vyd. 1. Praha: Grada Publishing, 2010, ISBN 978-802-4717-081.
- [7] LIU, Derong a Anthony N MICHEL. Dynamical systems with saturation nonlinearities: analysis and design. New York: Springer-Verlag, c1994. ISBN 03-871-9888-1.
- [8] MCDERMOTT, Joseph O'Connor. The art of systems thinking: essential skills for creativity and problem solving. London: Thorsons. ISBN 978-072-2534-427.
- [9] MEADOWS, Donella H a Diana WRIGHT. Thinking in systems: a primer. White River Junction, Vt.: Chelsea Green Pub., ISBN 16-035-8055-7.
- [10] MILDEOVÁ, Stanislava. Systémová dynamika: tvorba modelu. Vyd. 1. Praha: Oeconomica, 2011. ISBN 978-80-245-1842-8.
- [11] MILDEOVÁ, Stanislava, DALIHOD, Martin, EXNAROVÁ, Anna. Mental Shift Towards Systems Thinking Skills in Computer Science. Journal on Efficiency and Responsibility in Education and Science [online], 2012, roč. 5, č. 1, s. 25–35. ISSN 1803-1617. URL: http://www.eriesjournal.com/_papers/article_165.pdf.
- [12] MILDEOVÁ, Stanislava, DALIHOD, Martin, KRÁL, Miroslav. System Thinking in Informatics: a Case of COBIT. In: CONFENIS – 2013. Praha, 11.09.2013 – 13.09.2013. Linz: Trauner Verlag, 2013, s. 287–296. ISBN 978-3-99033-081-4.
- [13] MILDEOVÁ, Stanislava, DALIHOD, Martin. Validita modelu & realita systému. In: Systémové přístupy 2010 [CD-ROM]. Praha, 11.11.2010. Praha: Oeconomica, 2010, s. 83–90. ISBN 978-80-245-1728-5.
- [14] MOLNÁR, Zdeněk. Competitive intelligence, aneb, jak získat konkurenční výhodu. Vyd. 1. V Praze: Oeconomica, 2012. ISBN 978-80-245-1908-1.
- [15] STERMAN, John. Business dynamics: systems thinking and modeling for a complex world. Boston: Irwin/McGraw-Hill, c2000, ISBN 00-723-1135-5.
- [16] STRÍŽOVÁ, Vlasta. Presentace informací a komunikace. Vyd. 1. V Praze: Oeconomica, ISBN 978-802-4517-148.

- [17] ŠTECHA, Jan a Vladimír HAVLENA. Teorie dynamických systémů /: přednášky. Vyd. 2. Praha: Vydavatelství ČVUT, 1999. ISBN 80-010-1971-3.
- [18] ŠUSTA, Marek. Cvičení ze systémové dynamiky. Vyd. 1. Praha: Oeconomica, 2004. ISBN 80-245-0780-3.

JEL Classification: M150

Summary

System dynamics models usage when solving management communications issues

This paper deals with the application of system dynamics models in managerial practice. In today's world, where management and economic organizations already perceives the growing importance of systems thinking to solve problems, which are associated with the increasing complexity of information systems of the organization. For this time being there is no change of the linear paradigm that is firmly associated within idiosyncratic model of individual managers. This article aims to draw attention to the possibility of sharing mental model towards correcting the idiosyncratic model. Model, which is part of the managerial aspects of communications, and which is taken into account in managerial decision making.

For the purposes of article three issues were selected, covering three major areas of managerial communication. Specifically, the area of Vision, Project Management and Audit.

To his decision one nowadays very often uses the so-called Decision Support Systems (DSS), which leads to a confrontation of mental models in the form of partial solutions through information technology. Information technology enables questioning of the experiment with the real form of mental models. In practice, these models can be used as special or universal (general) – that help their users (managers) in the implementation of the management and decision-making activities. These model-based systems, to be fully functional, are not using the whole capacity of large enterprise databases. DSS are used is the very model creation itself, simulation, and to help analyze leverage points between variables (these are selected as a sample from the database - initially for the purpose of verification of the model and then for making predictions). System dynamics models (whose aim is to solve selected problems of managerial communication) are being described in the article as a draft form of decision support systems.

The article should contribute to the professional multidisciplinary discussion and provide some suggestions for consideration to the question: What competitive advantage can DSS, described in the article, bring to the current environment of economic organization, when used in managerial communication?

Keywords: System Dynamics, Decision Support Systems, mental models, system thinking, system, management communication, competitive advantage.

Studium vlivu behaviorálních a sociálních faktorů v organizačních procesech řízení kvality softwaru a dodávek informačních systémů

Michal Doležel

michal.dolezel@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: doc. Ing. Norbert Žid, CSc., (norbert.zid@vse.cz)

Abstrakt: Cílem tohoto příspěvku je metodologické vymezení výzkumu, který se zaměřuje na spojení konceptů řízení kvality softwaru (s důrazem na testování) a řízení outsourcingu v oblasti dodávek informačních systémů. Cílem výzkumu je návrh procesního rámce pro řízení těchto aktivit v byznys organizacích, a zároveň studium sociálních a behaviorálních faktorů v tomto kontextu. Příspěvek stručně pojednává o střetu perspektiv podnikové/byznys informatiky s disciplínou softwarového inženýrství, která je „tradičním domovem“ konceptu řízení kvality softwaru. Zároveň otevírá i otázku metodologických odlišností v pojmání vědeckých disciplín (německé) podnikové/byznys informatiky a (anglosaského) konceptu informačních systémů. Základní tezí příspěvku přitom je, že software lze vnímat jako podstatnou součást informačních systémů v byznys organizacích, avšak nikoliv jako součást jedinou. Příspěvek rovněž vysvětluje motivy, které autora vedou k nutnosti začlenit výzkum sociálních a behaviorálních faktorů do téhož výzkumného projektu a stručně představuje inovativní metodu výzkumu, která je pro tyto účely úspěšně použitelná. Jde o metodu akčního návrhového výzkumu.

Klíčová slova: vývoj informačních systémů, řízení kvality softwaru, řízení dodavatelů, outsourcing, akční návrhový výzkum.

Úvod

Ačkoliv na informačních technologiích (IT) již dnes „nezáleží“ (Carr 2003), na informacích, a tedy i informačních systémech (IS) v organizacích, záleží enormně: „IS a IT jsou dvě odlišné věci“ (Avison & Elliot 2006, s. 6). Organizace totiž využívají IS k bezprostřednímu uspokojování svých komplexních informačních potřeb, přičemž forma a obsah IS jsou pro ně klíčové. Problematika vývoje a dodávek počítačových IS formou outsourcingu není v této souvislosti nová. Avšak snad více než u jiných informatických oblastí se u ní projevuje fakt, že vývoj zakázkového softwaru nikdy nebude informatickou komoditou, jakou je třeba nákup „holých“ informačních technologií či „konfekčního“ programového vybavení.

Software (a následně i jeho kvalitu) lze přitom vnímat jako podstatnou součást IS v organizacích, avšak nikoliv jako součást jedinou (Von Hellens 1997). Kvalita softwaru, její řízení a zlepšování softwarových procesů (SPI) je přitom koncept, který je dlouhodobě etablovaný především v odvětví softwarového inženýrství (Sommerville 2007) a nezapře svou inspiraci oblastí tradiční průmyslové výroby a především filosofií Total Quality Managementu (Hansen et al. 2004). Softwarové organizace tyto principy tradičně používají ke konceptuálnímu uchopení organizačních a procesních otázek v souvislosti s uspokojováním potřeb zákazníka a produkcí jakostních a předvídatelných výstupů - softwaru. Řízení kvality softwaru je však možno vnímat i jako nezbytnou součást metodik vývoje IS (Buchalcegová 2009), popř. jako součást oblasti IT governance (Svatá 2012) či konceptu řízení podnikové informatiky (Voříšek & Pour 2012). Určitým způsobem zde totiž jde i o střet perspektiv, ze které badatel na tuto oblast nazírá (Doležel 2013a). Z hlediska řízení podnikové informatiky však nebylo u nás ani ve světě dosud tradicí tyto koncepty příliš spojovat (zejm. v patřičné úrovni detailu a s důrazem na řízení testování IS vyvíjených dodavatelským způsobem).

Jde však především o to, že využití zmíněných konceptů řízení kvality v prostředí outsourcingu vyvolává nejednu těžkost. Je možno klást si např. tyto otázky:

- Lze na software obdobně aplikovat principy tradiční výroby a subdodávek?
- Má odběratel etické právo průběžně kontrolovat a dozorovat práci dodavatele, „dívat se do kuchyně již během vaření“ s cílem preventivně intervenovat?
- Či je případný následný mimořádný audit IS z jeho strany dostačujícím prostředkem?

- Do jaké míry jsou obě varianty pro něj ekonomické?
- Jakou roli zde hraje dlouhodobá vzájemná důvěra a ekonomické výhody pro obě strany?
- Jaké organizační role za stranu odběratele by na těchto procesech měly participovat a jak nastavit jejich organizační odpovědnosti?
- A jaké důsledky přináší tlak na minimální cenu dodávky ze strany odběratele, resp. jak najít optimální poměr cena/kvalita?
- Má vliv i konkrétní kulturní kontext odběratelské organizace?
- Jakou roli prisuzovat dalším sociálním a behaviorálním faktorům?

Podobné otázky jsou vyvolány především potřebami praxe, ale „správné a ověřené“ odpovědi na ně většinou nejsou známy. Cílem tohoto příspěvku je představit metodologická východiska výzkumu, ve kterém se zaměřuji na problematiku dodávek IS s důrazem na koncept řízení kvality softwaru (zejm. testování) a související tzv. měkké faktory tak, jak byly stručně představeny výše. Cílem výzkumu je přitom odpovědět na část položených otázek¹. Z důvodu limitu délky příspěvku zde nemám možnost představit již konkrétní dosažené výzkumné výsledky (zejm. výsledky návrhového výzkumu) a získaná data – tyto byly či budou předmětem jiných příspěvků. Naopak si kladu za cíl objasnit použitou metodologii vědecké práce, která není v českém prostředí zcela typická a vychází z některých u nás méně tradičních přístupů využívaných dosud zejm. v anglosaském prostředí.

Příspěvek je strukturován následujícím způsobem: nejprve je představeno zaměření výzkumu, konkrétně je diskutováno oborové zařazení a způsob začlenění výzkumu sociálních a behaviorálních faktorů do původně čistě informaticky orientovaného výzkumu. Následně se přesouvám k metodologickému rámci výzkumu, tedy filosofickým, paradigmatickým a epistemologickým východiskům výzkumu, jakož i konkrétním použitým výzkumným metodám a technikám. V závěru jsou stručně shrnuty klíčové poznatky.

Zaměření výzkumu

Tato část příspěvku se věnuje zaměření výzkumu, kterým je zde myšleno především jeho oborové zařazení a vyjasnění rozdílů v chápání jednotlivých informatických disciplín z pohledu badatelských tradic v České republice a ve světě. S tím souvisí i otázka volby vhodných odborných zdrojů. Dále tato část ukotvuje propojení převážně technického tématu s oblastí společenských věd.

Vymezení výzkumného pole

Výzkum je orientován transdisciplinárně, avšak jeho těžiště leží ve dvou aplikovaných informatických disciplínách² (srov. ACM/AIS/IEEE-CS 2005) – softwarovém inženýrství (SE) a informačních systémech (IS). Označení „IS“ však u nás dosud bohužel nemá, na rozdíl od západního světa, na akademické půdě silnější tradici, proto přistupuji níže k jeho vymezení. Hlavní důvod tohoto stavu (tzn. roztržičnosti terminologie pro označení dílčích „computing“ oborů) spatřuji v historických souvislostech. Např. Vlasák (2009) připomíná, že použití termínu „informatika“ pro počítačovou vědu (computer science) v jejím angloamerickém chápání lze z historické perspektivy vnímat jako aktivní vzdor Francouzů proti Američanům, který – po rozšíření pojmu do dalších evropských zemí – do značné míry vyústil v terminologické „zmatení jazyků“ na kontinentu. V té době se u nás totiž, pod tíhou Sovětského svazu, již používal pojem informatika především pro obor zabývající se „komunikací informací *ve vědě*“ (ibid., nepaginováno). Obdobně, obsah anglického slova „Informatics“ je mnohem širší než pouze „počítačová věda“. V tamní tradici jde totiž o nadřazenou kategorii k pojmům, jakými jsou např. informační systémy, informační management, podniková informatika, sociální informatika, informační věda (!) aj. (Beynon-Davies 2007; Vlasák 2009).

¹ Nejde o exaktní „research questions“ - tyto otázky mají pouze ilustrativní a srozumitelnou formou vymezit rozsáhlou problematiku mého výzkumu.

² Pojem „disciplína“ používám ve smyslu existujících uzancí na mezinárodním výzkumném poli, kdežto pojmem „obor“ naznačuji stav v České republice z hlediska institucionálních akreditací studijních oborů a programů. Proto chápu např. počítačovou vědu (computer science) jako disciplínu, kdežto informatiku jako obor. Je však zřejmé, že mezi obsahovými náplněmi je větší či menší překryv. Podobný stav je v aplikované informatice, která je mým „akademickým domovem“.

Označení IS je přitom ve vědeckém světě chápáno, pomínou-li pregraduální vzdělání a firemní praxi, především jako název deskriptivně zaměřené vědecké disciplíny (Baskerville & Myers 2002; Avison & Elliot 2006; Avgerou 2000) operující na výzkumném poli, které sdílí s více intervenčně zaměřeným, vědecko-prakticky pojímaným, informačním managementem, popř. podnikovou/byznys informatikou (Doležel 2013b; Buhl & Lehnert 2012). Světoví autoři však zároveň běžně disciplínu IS a informační management považují za synonyma (např. Beynon-Davies 2007), popř. chápou „management informačních systémů“ jako jednu z náplní širší disciplíny IS (Avgerou 2000). Ať tak či tak, většina lokálních badatelů se však patrně shodne, že domovem technologicky orientované „computing“ disciplíny informační systémy (nikoliv informační vědy, která má svůj domov na humanitně zaměřených pracovištích) jsou a mají být české akademické obory „aplikovaná informatika“, „management“ a obory příbuzné, mající co do činění s nasazováním informačních technologií v prostředí byznysu - podobně jako je tomu u konceptu podnikové informatiky a informačního managementu.

Ve své práci se obsahově hlásím ke konceptu, který Morrisonová a George (1995) nazývají „[výzkumný] polel ustaveným kombinací [disciplín] MIS³ a softwarové inženýrství ... [a] zahrnujícím všechny aspekty vývoje informačních systémů z perspektivy organizace“. Nadřazená perspektiva IS, prolínající se celým výzkumem jako jeho hlavní ukotvení, je v mém podání zdůrazněna především faktem, že se zaměřuji převážně na organizace, jejichž hlavní podnikatelskou činností není tvorba softwaru (jde tedy o tzv. „byznys organizace“). V současné době je u těchto organizací častá snaha o outsourcing veškerých činností, které nelze považovat za jejich „kmenový byznys“. Vývoj informačních systémů pak není výjimkou. V rámci svého výzkumu však vycházím z předpokladu, že ani tento typ organizací nemůže rezignovat na snahy držet dodávky svých informačních systémů alespoň částečně ve vlastních rukou a na problematiku nenazírat pouze jako na „black-box“. Proto volím (v návrhové části výzkumu) přístup, který kombinuje znalosti a přístupy obou disciplín (tedy IS a SE) a onen abstraktní „black-box“ částečně rozkrývá. Můj návrhový přístup lze tedy charakterizovat inženýrskou metaforou jako „grey-box“, avšak nikoliv jako „white-box“, který by analogicky odpovídal čistému přístupu SE a zajímal by ho každý procesní a technický detail vývoje softwarové aplikace. Z organizační perspektivy disciplíny IS tedy zkoumám a řeším část inženýrských a procesních problémů disciplíny SE (srov. Avgerou 2000). Jde zejm. o organizační aspekty a tzv. měkké faktory související s vývojem informačních systémů a jejich dodávkami.

Způsob začlenění výzkumu sociálních a behaviorálních aspektů

Výše popsáný přístup mi tak (v pozdější fázi výzkumu) umožňuje oprostit se od většiny čistě technických a rovněž i části procesních problémů, a v této ověřovací části se zaměřit především na problematiku zkoumání lidských aspektů v rámci vývoje a dodávek informačních systémů. Tato oblast není dosud v literatuře dostatečně zmapována a její výzkum lze označit jako určitý nastupující světový trend, který začíná měnit i samotnou podobu původně „čistě technického“ softwarového inženýrství. Jak např. upozorňuje populárně-vědecký časopis *American Scientist*, „softwarové inženýrství se nyní nachází na zásadní křižovatce srovnatelné s úsvitem medicíny založené na důkazech (evidence based medicine), kdy zdravotnická komunita začala zkoumat své vlastní praktiky a oddělovat ty které opravdu fungovaly od těch, které byly jen mýty“ (Wilson & Aranda 2011, s. 466). V souvislosti se zájmem o sociální dopady vývoje softwaru/IS je vhodné dodat, že samotná vědecko-technická identita softwarového inženýrství je pro mnohé pracovníky praxe ošidná a falešná – podle nich návrh a tvorba softwaru nikdy nebude totéž jako projektování mostu či výpočet rezonanční frekvence oscilátoru. „Naopak, programování je kvalifikované řemeslo, daleko bližší architektuře [než stavebnímu či elektrotechnickému inženýrství], což činí lidský element důležitým (někdo by řekl zásadním) předmětem studia“ (Wilson & Aranda 2011, s. 466).

Disciplínami, které mi tedy umožňují zaměřit se právě na tuto problematiku z perspektivy oboru aplikované informatiky, je jednak v posledních letech bouřlivě nastupující empirické softwarové inženýrství (ESE), ale především již zmíněná disciplína IS. Ta později uvedená již tradičně, několik desetiletí silně těží ze sociálně-vědního světa, konkrétně ze sociologie, psychologie a sociální/kulturní antropologie

³ Zkratka MIS, Management Information Systems, je označením používaným zhruba do roku 2000. Nyní se disciplína již běžně označuje jen jako IS.

(Avison & Elliot 2006). Její vědecko-výzkumnou identitu lze výstižně charakterizovat takto: „Zatímco informatik či systémový inženýr se bude dívat na technologii, která stojí ve středu jeho zájmu, výzkumník v oblasti IS se dívá stranou počítačů – zajímají ho lidé a organizace.“ (str. 7–8). Je běžné, že badatelé v této disciplíně těží rovněž z teorie organizační vědy/organizačního chování (Orlikowski & Barley 2001), v naší vysokoškolské tradici dosud spíše nazývané „teorie managementu“ či „teorie podnikového řízení“.

Ačkoliv přesto mohou zaznít námitky, že výzkum sociálních aspektů nepatří do oboru (aplikované) informatiky, tvrdím, že tyto námitky by pomíjely minimálně dva hlavní argumenty proč tomu tak není: (1) znalost těchto zákonitostí má obrovský potenciální dopad na podnikovou praxi – a současný stav v této oblasti je přitom možno označit jako nepochybné „tápání ve tmě“, (2) nelze příliš očekávat, že tuto oblast výzkumu si za své vezmou sociální vědy. Je pro ně totiž příliš technická, svým způsobem i málo zajímavá. Mikrokosmos organizací a jejich IT oddělení je totiž nepochybně mnohem méně atraktivní než třeba marginalizované skupiny městské společnosti, kupř. bezdomovci (srov. např. Hejnal 2013), a navíc vyžaduje rozsáhlé oborové znalosti a předchozí zkušenosti s technickou problematikou. Druhý zmíněný argument však nikterak nebrání tomu, aby badatelé v této oblasti přijali za své kvalitativní metody výzkumu, jakož i sociální teorie, jejichž původním domovem jsou právě sociální vědy. Popř. podobně jako já, aplikovali hybridní přístup založený na kombinaci organizační intervence (východiskem je „klasická“ metoda návrhového výzkumu) a následného intervenčně-popisného výzkumu dopadu (např. metodou akčního výzkumu aj.), což je přístup charakteristický např. pro akademickou disciplínu nazývanou organizační změna a rozvoj (Marrewijk et al. 2010).

Současné diskuze na světovém vědeckém poli svázaném s disciplínou IS jsou poznamenány především již naznačeným pnutím mezi chápáním teoretičtější zaměřené disciplíny IS v anglosaské tradici proti pragmatickému pojetí byznys informatiky v tradici německy mluvících zemí (Buhl et al. 2012; Doležel 2013b), tedy bojem mezi „sociálním konceptem“ a „inženýrským přístupem“. Já hledám polohu podle mého názoru optimální, nacházející se mezi oběma polohami extrémními. Na jednu stranu si uvědomuji, že současný angloamerický přístup (ač vysoce rigorózní a vědecky sofistikovaný) je zřídka přímo užitečný pro praxi, na druhou stranu je mi však na základě mých profesních zkušeností zřejmé, že ani ryze technokratický přístup mající blízko snad až k „sociálnímu inženýrství“ nevede k požadovaným reálným výsledkům v praxi zejm. kvůli tomu, že často zcela opomíjí lidský faktor (srov. Molnár et al. 2012). Organizace není stroj – je sice možno ji inženýrskými metodami navrhnout, ale „vyrobit“ (tzn. implementovat navržené postupy v praxi) již takto přímočaře („inženýrsky“) nelze (Pentland & Feldman 2008). Ukázkovým příkladem je např. fatální selhání informaticko-inženýrského organizačního konceptu BPR (Business Process Reengineering) v devadesátých letech 20. stol. v podnikové praxi (Chytil 2010). Z poznatků získaných během mého výzkumu v této souvislosti především vyplývá, že vždy je nutno respektovat i kontext - tzn. zajímat se o sociální a behaviorální faktory v organizaci, pro kterou z návrhu „vyrábím“ výsledný produkt – k implementaci určený organizační artefakt (Hevner & Chatterjee 2010). Můj výzkumný přístup tak hledá tolik potřebnou rovnováhu mezi užitečností pro praxi a teoretickou zvučností, přičemž je vedený kvalitativní metodologií akčního výzkumu využívající rovněž návrhový výzkum (viz níže).

Dilemata syntetizujícího přístupu

Je třeba připustit, že studium „měkkých aspektů“ v organizacích přináší určité metodologické těžkosti. Především jde o jistý zakořeněný konflikt mezi psychologem na straně jedné, a sociologem a antropologem na straně druhé. Na tomto místě cítím potřebu zdůraznit, že se vymezuji vůči určité historické „nadvládě psychologie“ (včetně aplikovaného oboru psychologie práce a organizace) v této oblasti. Hlásím se k tradici informačních systémů a organizačního chování (studií) jako disciplín syntetizujících (eklektických) a pluralistických, tak jak jsou chápány především v anglosaské akademické tradici. Zde jsou jejich přirozeným domovem „business schools“, školy aplikované ekonomie, ekonomiky a managementu, včetně managementu IT. Ve většině případů tyto disciplíny nejsou provozovány na tradičních univerzitách. V tomto pojetí si tyto interdisciplinární obory (našemu chápání bližší spíše v označení „interdisciplinární výzkumná pole“) berou ze všech svých referenčních oborů různé mnohdy vzájemně konfliktní teorie, které následně aplikují a rozvíjejí v zajetí svých vlastních metod a akademických tradic.

Psychologii řízení a organizace v žádném případě nezavrhuji. Ačkoliv však oceňuji jedinečnou publikaci M. Nakonečného *Sociální psychologie organizace* (Nakonečný 2005) i další české monografie z pera psychologů (např. Bělohlávek 1996), beru pouze v úvahu i pohled kvalitativní sociologie a sociokulturní antropologie. Např. Nakonečný chápe problematiku lidských faktorů v organizacích především jako „aplikovanou sociální psychologii“. Já jsem si však zároveň vědom poznámky E. Scheina, amerického sociálního psychologa a později velmi významného organizačního teoretika na MIT Sloan School of Management (citace jeho prací lze nalézt i v českých publikacích (např. Lukášová 2010)) o určité omezenosti psychologického přístupu. Schein (1996) říká, že „většina ‚business schools‘, jež implementovala tento koncept [studia organizací], zaměstnala průmyslové a sociální psychology a pojmenovala ho jako ‚organizační chování‘ – nálepkou, se kterou jsem se nikdy nesmířil, protože ve mně vyvolává pochybnost jako nekonceptční protimluv“. Schein proto preferuje širší termín ‚organizační studia‘ (ibid.) a uzavírá na tuto dobu (devadesátá léta 20. stol.) převratnou myšlenkou, že „[organizační] psychologové nevěnovali dostatečnou pozornost sociologům a antropologům, jejichž [výzkumná] tradice spočívají v tom jít do terénu a dlouhodobě pozorovat určitý fenomén předtím, než se mu budou snažit porozumět“ (s. 231).

Taková podoba terénního výzkumu je nejčastěji nazývána etnografickou metodou a je založena na technice zúčastněného pozorování jako svém primárním metodologickém východisku. Své kořeny má především v sociální a kulturní (sociokulturní) antropologii. Sjednocující pohled této disciplíny a jeho důležitost pro svět organizací lze přitom shrnout např. takto: „Naše [antropologická] teorie je naše síla. Teorie managementu je poháněna sociologií a psychologií, a má *problémy překlenout propast mezi makroúrovni a mikroúrovni chování*. My máme dispozice propast překlenout. Vidíme strukturu. Vidíme cesty k porozumění chování jednotlivců jako části chování celku. Zatímco naši užitečnost pravděpodobně lidé v organizacích budou vidět spíše v oblasti porozumění cizím kulturám, my jim můžeme být úplně stejně užiteční v porozumění jejich organizacím.“ (Jordan 1994, s. 1156, *zvýraz. doplněno*). Zkoumání byznys organizací z antropologické perspektivy se přitom obšírněji věnují jinde (Doležel 2013b).

Z pohledu české sociálně-organizační literatury lze tento koncept označit nálepkou „*kulturologická koncepce organizace*“, přičemž explicitně předpokládám, že z hlediska obsahové náplně mého výzkumu se „kultura organizace a její vliv ... projevuje také ve střetu individuálních přání a kulturně požadovaných vzorů při jejich naplňování“ (Nový et al. 2006, s. 42). V té souvislosti v zásadě odmítám „klasický“ Hofstedeho model dimenzí národní kultury (Lukášová 2010). Národní a organizační kulturu pokládám v souladu s Myersem a Tanem (2002) za natolik komplexní fenomén, že jeho prosté kvantifikování do pěti (šesti) numericky exaktně vyjádřených dimenzí považuji za problematické (srov. Cater-Steel & Toleman 2008).

Mimo Scheinova konceptu organizační kultury mapujícího i souvislost s národní a profesní kulturou (Schein 2010) se hlásím k dílu Henryho Mintzberga, známého profesora managementu hlásajícího nutnost propojení teorie a praxe a kritizujícího současný stav amerického MBA studia. Mintzberg, před více než 40ti lety, přispěl svou disertací (Mintzberg 1968) k položení základů využití etnografické metody v organizačních studiích, tehdy ještě pod citelným vlivem sociální psychologie nutícím ho používat „strukturované pozorování“ (Czarniawska 2012). Dále se svým výzkumem hlásím k práci G. Kundy (2009), který etnografickou metodou obšírně studoval IT organizace a jejich kulturu. Z hlediska použitých metod se tedy orientuji na koncept etnografického výzkumu, který je „antropologickou nebo sociologickou [výzkumnou] aktivitou *mnohem více než aktivitou psychologickou*“ (Watson 2011, s. 213, *zvýraz. doplněno*). Avšak tento koncept přizpůsobuji potřebám oblasti výzkumu byznys informatiky/informačních systémů v organizacích (využívám ho uvnitř intervenční metody akčního výzkumu – viz dále), a vybírám si z něj především jeho hlavní techniku – zúčastněné pozorování či „pozorovací zúčastnění“ (Moeran 2009).

Takový přístup je poměrně inovativní a v českém prostředí může být pokládán snad až za kacířský i vzhledem k tomu, že ani jedna ze zmíněných syntetizujících disciplín studia organizací a IS (v jejich angloamerické podobě) u nás není dosud ustavena jako obor samostatný a takzvaně tradiční. Navíc, jde vesměs o velmi nové poznatky, které je možno nalézt pouze v originální cizojazyčné literatuře z posledních cca. 10-20 let. Přenášení těchto poznatků do českého prostředí tak bude nutně vyvolávat otázky týkajících se legitimity takového přístupu, protože jde proti zavedeným lokálním normám a tradicím.

Např. ne každý bude zpočátku připraven přistoupit na myšlenku, že v anglosaském kontextu je současná teorie managementu vnímaná především jako syntetizující sociální věda (!) nazývaná jako „organizační chování“ či „organizační studia“. Tato má pak mnohem blíže ostatním sociálním vědám, než třeba ke staršímu (a dnes již veskrze překonanému) konceptu praktického nasazování „management science“ (operační výzkum) jako reprezentantu linie tzv. „tvrdé vědy“ (srov. Molnár et al. 2012). Podobně nezáviděníhodnou pozici lze kontextu České republiky přisuzovat i kvalitativním výzkumným metodám, které jsou typicky především se sociálně-vědní oblastí (odkud pocházejí) dodnes téměř výlučně spojovány. Slovy geografů řešících podobné výzvy ve svém oboru: „Kvalitativní výzkum ... není v postsocialistických zemích příliš rozvíjen. Lze dokonce pozorovat, že jeho absence či velmi pomalý rozvoj není nahodilou strukturní vlastností postsocialistických společností. Je možné to interpretovat především na pozadí skutečnosti, že kvalitativní metodologie čerpá v základu z filozofických směrů, které byly před rokem 1989 v našich podmínkách nepřijatelné, a autorů, kteří, když nebyli přímo `zakázáni`, tak byli alespoň výrazně dezinterpretováni“ (Osman et al. 2011, s. 22).

K využívání poznatků sociálních věd se u nás dosud jako součást aplikované informatiky hlásil, především informační management (IM), který je však velmi blízký konceptu IS a nezanedbatelně se s ním (v rovině teorie) překrývá, jak již bylo naznačeno výše. Koncept IM je ze své podstaty definován jako technicko-společenské transdisciplinární pole orientované směrem k praxi, spojující informatiku a management při zohlednění toho, že v lokální tradici klade velký důraz na systémový přístup (Vodáček & Rosický 1997). Na systémový přístup klade důraz i disciplína IS, avšak v rámci její tradice získaly historicky velký vliv i další přístupy, např. koncept označovaný jako organizační realismus, Giddensova teorie strukturace pocházející ze sociologie a rovněž kritická teorie (Avgerou 2000). Je si totiž třeba uvědomit, že pokud připustíme myšlenku „teorie managementu jako sociální vědy“, musíme následně připustit i to, že u IM (a IS) půjde v oblasti teorie o *boj mezi „měkkou“ teorií společenských věd a „tvrdou“ teorií informatiky*.

V této souvislosti může rovněž zaznít námitka týkající se přílišné „teoretičnosti“ mého výzkumu. Tuto však svým intervenčně-popisným přístupem eliminuji. Uvědomuji si, že z pohledu technika se skutečně může aplikace vědecké teorie psychologie, sociologie a antropologie do oblasti informatiky jevit kontroverzně. Lze se např. ptát: „jak je možné, že na totožný organizační fenomén je možno aplikovat různé sociálně-vědní teorie, které nejsou vzájemně ‚kompatibilní‘ a náleží dle svého zaměření často právě jen do jednoho z těchto oborů?“. Je si však nutno uvědomit, že podobný názorový pluralismus je sociálním vědám vlastní a nelze ho vnímat „inženýrskou binární logikou“ jako „dobré“ či „špatné“ řešení problému, tedy „dobrou“ či „špatnou“ teorii.

Podobně si je nutno uvědomit, že výchozím přesvědčením badatelů v disciplíně IS (v jejím anglosaském pojetí) je názor, že společenskovední teorie z pohledu jejich globální platnosti a aplikovatelnosti na společnost nelze izolovat od mikrosvěta organizací. Tyto mikrosvěty jsou totiž rovněž součástí globální společenské reality – minimálně jimi prostupují „titíž lidé“, kteří hrají hlavní úlohu v „globálních“ společenskovedních teoriích. Tato myšlenka je hlavním důvodem, proč se moderní disciplíny zabývající se studiem organizací zajímají o teoretické dědictví „tradičních a globálních“ společenských věd (zejm. psychologie, sociologie a antropologie) a navazují na ně. A informační systémy obecně (tzn. nejen ty počítačové) – chápány jako klíčové elementy těchto organizačních mikrosvětů – jsou tedy i klíčovým předmětem studia vědecké disciplíny IS, která zkoumá technické a sociální problémy související s vývojem, provozem a řízením informačních systémů v kontextu organizací (Walsham 1993).

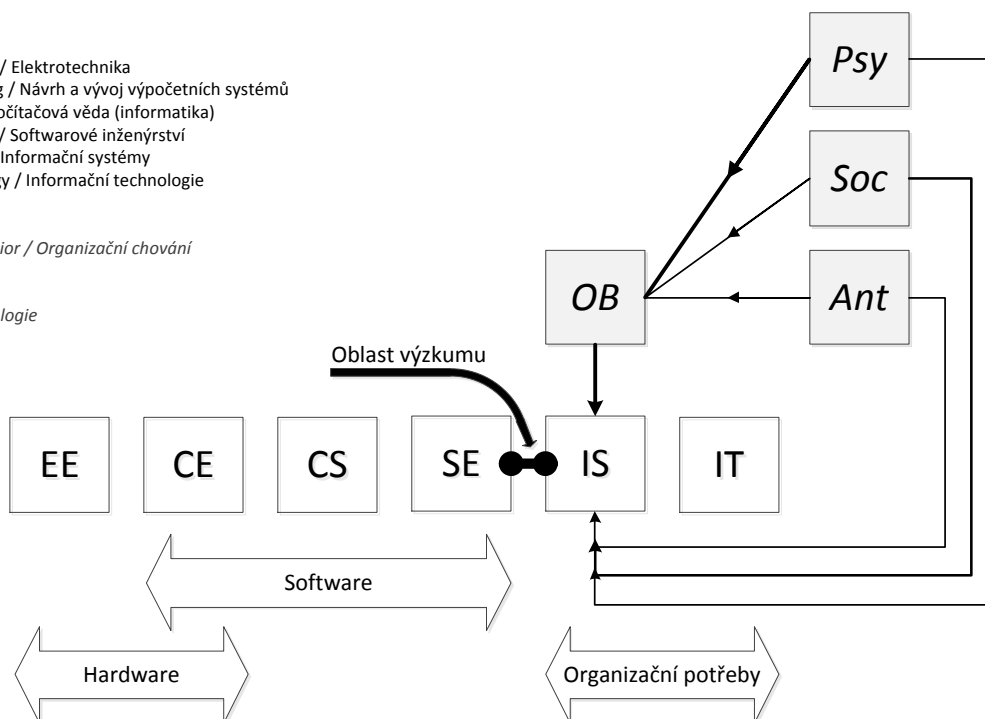
Koncept využití poznatků sociálních věd a transdisciplinarita mého výzkumu jsou demonstrovány obrázkem 1. V obrázku jsou v pravé části silnější linkou zdůrazněny majoritní vlivy referenčních disciplín.

Legenda:**Technické disciplíny:**

EE – Electrical Engineering / Elektrotechnika
 CE – Computer Engineering / Návrh a vývoj výpočetních systémů
 CS – Computer Science / Počítačová věda (informatika)
 SE – Software Engineering / Softwarové inženýrství
 IS – Information Systems / Informační systémy
 IT – Information Technology / Informační technologie

Společenské vědy:

OB – Organizational Behavior / Organizační chování
 Psy – Psychologie
 Soc – Sociologie
 Ant – Sociokulturní antropologie



Obrázek 6: Předmět výzkumu a vztah disciplíny IS ke společenským vědám

Zdroj: autor, inspirováno (ACM/AIS/IEEE-CS 2005)

Metodologie výzkumu

Tato část příspěvku popisuje již konkrétní filosofická, paradigmatická a epistemologická východiska, která mají bezprostřední dopad na způsob provádění výzkumu. Rovněž zde pojednávám o použitých výzkumných metodách.

Filosofická a epistemologická východiska výzkumu

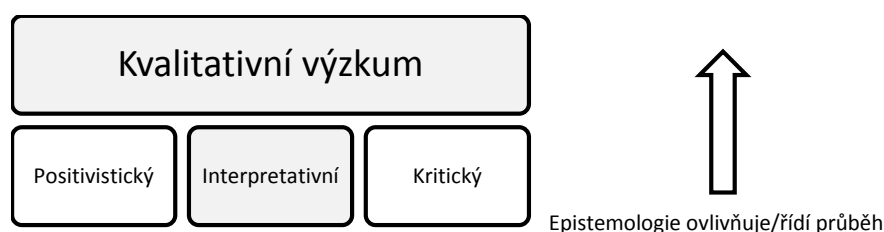
V předchozích částech jsem vymezil pozici svého výzkumu v rámci informatických disciplín. Neméně závažným rozhodnutím při studiu sociálních aspektů v organizacích je ovšem otázka volby epistemologie. Jde vlastně o volbu přístupu ke studiu sociální reality, z nichž dosud tradičním a snad i nejfrekventovanějším je především pozitivismus, historicky spojovaný s osobností filosofa A. Comta. Základní myšlenka pozitivismu spočívá ve víře, že „sociální realita existuje externě (tj. nezávisle na člověku) a že její vlastnosti lze zkoumat a měřit prostřednictvím tzv. objektivních metod“ (Pavlica 2000, s. 27). A dále, „že realita ve které žijeme je daná, svoji podstatou neměnná a uniformní“ (s. 28). Takový postoj pak ve výzkumu akcentuje především kvantitativní metody jako nástroje tzv. „objektivního měření“. Avšak zejména ve světě se stává terčem stále sílící kritiky (např. Davison & Martinsons 2011) fakt, že pozitivistický výzkum v této podobě neumožňuje zachytit motivy jednání a „skutečné světy“ jednotlivých účastníků výzkumu, tzn. konkrétních a živoucích osob, jakož i porozumět složitějším a reálným organizačním jevům.

Z tohoto důvodu se při studiu složitějších organizačních jevů svět více a více obrací ke kvalitativním metodám spojeným často s interpretativní epistemologií. Tento přístup lze ilustrativně přiblížit v souladu se známou tezí formulovanou Geertzem: „Domnívá se, společně s Maxem Weberem, že člověk je zvíře zavěšené do pavučiny významů, kterou si samo upředlo, považují kulturu za tyto pavučiny a její analýzu tudíž nikoliv za experimentální vědu pátrající po zákonu, nýbrž za vědu interpretativní, pátrající po významu.“ (Geertz 2000, s. 15, zvýraz. doplněno). Výstižně byla tato myšlenka zachycena i v souvislosti s jedním konkrétním kvalitativním disertačním výzkumem organizační kultury prováděným v rámci oboru sociologie. Sociologie je v lokálním pojetí (patrně hned po ekonomii) jedním z „nejtvrdších“ společenskovědních oborů vůbec, tradičně akcentujícím především kvantitativní metody, statistické vzorky a vůbec „tvrdá data“ - česká tradice kvalitativní sociologie totiž není příliš dlouhá a

silná (zjevná paralela k lokálním disciplínám studujícím organizace). Mj. proto autorka ve své disertační práci obhazuje volbu interpretativní epistemologie touto rétorikou: „Nejde mi o popis nějaké ‚objektivní‘ pravdy a toho, jak to v ... [organizacích] ‚dopřavdy‘ chodí, ale místo toho se snažím uchopit významy a světy svých informátorů tak, jak je sami konstruují, nahlíží, vnímají a interpretují.“ (Hrešanová 2008, s. 36).

Podobně jako se k těmto tezím hlásí někteří kvalitativní sociologové a organizační teoretikové tíhnoucí především k antropologii, resp. etnografii organizací (srov. Doležel 2013b), pokládají Geertzovu Interpretaci kultur za svou „bibli“ (Bate 1997, s. 1197), navazují na ni i badatelé v disciplíně IS. Významným představitelem tohoto směru je v této disciplíně Geoff Walsham, profesor IS na Judge Business School na univerzitě v Cambridge, který svou knihou *Interpretace informačních systémů v organizacích* (Walsham 1993) položil v devadesátých letech 20. stol. základy této výzkumné filosofie. Následovaly další, dnes velmi citované články (Walsham 1995; Walsham 2006), které ukotvují postavení interpretativní epistemologie v rámci disciplíny IS. Ačkoliv v devadesátých letech 20. stol. se k interpretativní perspektivě v disciplíně IS hlásilo pouze 12% článků zveřejněných v EJIS, špičkovém evropském vědeckém časopise zaměřeném na informační systémy, a 18% ve vybraných amerických špičkových IS časopisech, po roce 2000 se tento trend markantně zesiluje právě ve prospěch tohoto typu výzkumu. Tímto směrem disciplína IS v současné době zjevně spěje (Walsham 2006).

Obrázek 2 schematicky znázorňuje úlohu epistemologie při provádění kvalitativního výzkumu. Zvolená perspektiva relevantní mému výzkumu je zvýrazněna.



Obrázek 2: Výchozí filosofické předpoklady behaviorálního výzkumu Zdroj: (Myers 2009)

Použité metody kvalitativního a návrhového výzkumu

Výše uvedené skutečnosti jsou sice zásadní, avšak jen velmi málo vypovídají o konkrétních použitých výzkumných metodách - mají však vliv především na širší metodologické ukotvení a povšechná východiska výzkumu. V souvislosti s konkrétním postupem výzkumu v oblasti IS je však nutno především vyřešit otázku volby konkrétní výzkumné metody a s tím související otázku výzkumných technik (např. Pavlica 2000; Myers 2009).

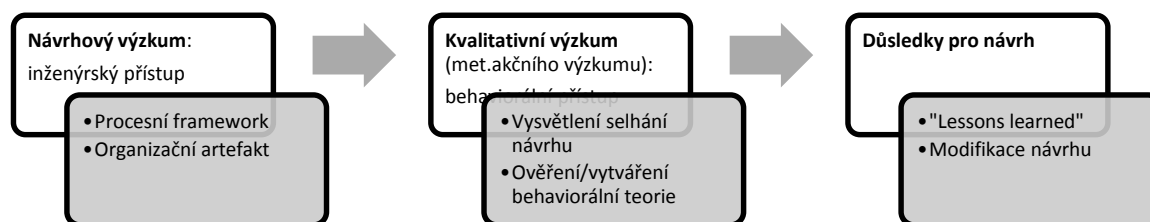
Do metod kvalitativního výzkumu v disciplíně IS se přitom řadí níže uvedené. Jejich použití má typicky induktivní charakter sloužící k vytváření teorií daleko častěji než k testování konkrétních hypotéz (jež by byly typicky znakem pozitivismu, jehož použití v rámci kvalitativního výzkumu není úplně typické). Podrobnou charakteristiku kvalitativních metod lze dohledat jinde, např. v Myersovi (2009). Ve stručnosti jde tedy o:

- Případovou studii (deskriptivní metoda využívající primárně techniky rozhovorů),
- Etnografický výzkum (deskriptivní metoda využívající primárně techniky zúčastněného pozorování, rozhovory jsou však pochopitelně rovněž její důležitou součástí),
- Zakotvenou teorii (specifická deskriptivní metoda, nová teorie jakoby „vyrůstá“ z dat),
- *Akční výzkum* (intervenčně-deskriptivní metoda nasazovaná v prakticky orientovaných projektech spolupráce badatelů a praxe, využívá různé techniky získávání primárních dat).

V oblasti podnikové/byznys informatiky (BISE) pak dosud jasně převládá návrhový výzkum spojený s tvorbou artefaktů (Österle et al. 2011), ačkoliv i všechny výše uvedené metody jsou v německé jazykové oblasti pokládány za zcela legitimní, byť řidčeji zastoupené a využívané (Wilde & Hess 2007).

Pro svůj výzkum jsem zvolil (na základě nutnosti překročit hranice běžných návrhových metod z důvodu pouze parciální úspěšnosti implementace navrženého artefaktu) inovativní kombinaci výzkumu návrhového a kvalitativního (v tomto případě konkrétně akčního výzkumu). Základní tezi použití této nové metody znázorňuje obrázek 3. Výzkum byl započat jako klasický návrhový výzkum produkující procesní rámec pro řízení kvality softwaru v dodávkách IS formou outsourcingu. V souvislosti s tím byla řešena i otázka návrhu organizačního uspořádání, a to včetně organizačního uspořádání dodavatelsko-odběratelských vztahů (na obrázku označeno jako „organizační artefakt“). Návrh byl úspěšný ve většině případů (projektů) na kterých byl ověřován, nicméně jeho implementace vykazovala nedostatky v případě jednoho softwarového programu (tzn. řetězce softwarových projektů) s řádově delší organizační historií než u zbývajících projektů. Vzhledem k tomu, že tato skutečnost jasně ukazovala na důležitost organizačního kontextu, který je *častým problémem při generalizaci výzkumných závěrů obdobných výzkumných projektů*, bylo toto impulzem pro zahájení druhé etapy výzkumu. V této etapě došlo k formálnímu přerámcování celého výzkumného projektu na kombinaci návrhového a akčního výzkumu. Předmětem akčního výzkumu přitom bylo především ověření artefaktu navrženého metodami typicky používanými v německé tradici disciplíny BISE, a to v intencích deskriptivně pojímané anglosaské disciplíny IS. To implikuje mj. nepominutelné zapojení sociálně-vědních teorií, které pomáhají osvětlovat příčiny nedostatků navrženého artefaktu, příp. důvody jeho parciálního selhání, a zdokonalit ho v budoucnu.

Kombinaci návrhového a akčního výzkumu lze pokládat za inovativní vědeckou metodu se značnou praktickou relevancí a zároveň nezanedbatelným teoretickým přínosem. Nasazování podobných postupů předznamenává ve vědecké komunitě IS/BISE postupné sbližování německé tradice výzkumu s tradicí anglosaskou, tedy se snahu o zvýšení vědecké prestiže podnikové/byznys informatiky (BISE) směrem k disciplíně IS, avšak při zachování relevance BISE pro praxi. Details mnou navrhovaného přístupu jsou popsány v samostatném příspěvku, který se momentálně nachází v recenzním řízení významné IS konference. V současné (počátek roku 2014) anglosaské časopisecké literatuře lze obdobný metodický postup dohledat pod označením „akční návrhový výzkum“ (action design research) (Sein et al. 2011). V odkazovaném případě jde však o jiný typ artefaktu než v případě mého výzkumu. Celkově je však nutno říci, že jde o velmi mladé postupy, u nichž vědecká terminologie není dosud ustálená, a uzance, terminologie a „best practise“ se teprve ustavují.



Obrázek 3: Navržený inovativní přístup k výzkumu (syntéza dvou typických přístupů) Zdroj: autor

Závěr

Řízení kvality softwaru v souvislosti externími dodávkami IS lze v době, kdy byznys společnosti začínají stále více a více přistupovat k částečnému či kompletnímu outsourcingu dodávek IS, považovat za klíčovou kompetenci IT oddělení těchto společností. Z hlediska informatické vědy tady však jde o střet minimálně dvou světových názorů a filosofí: na jedné misce vah jsou umístěna hlediska IT manažerů byznys organizací (perspektiva disciplíny IS/BISE), na druhou je nutno položit hlediska jednotlivých dodávajících softwarových společností a jejich manažerů softwarových projektů (perspektiva disciplíny SE). Výzkum v této oblasti si žádá nasazení inovativních vědeckých postupů, které jsou schopny překlenout tradiční a úzce chápáné koncepty této problematiky – na jedné straně zastoupené high-level „byznys“ konceptem IT governance, na druhé straně zastoupené detailistickým a „technicistním“ konceptem řízení kvality softwaru (Doležel 2013a). Do tohoto dilematu se navíc (jako další faktor) aktuálně promítá stále sílící střet německé podnikové/byznys informatiky BISE s anglosaskou disciplínou IS, který postupem času nepochybně silněji dolehne i k našim uším (Doležel 2013b).

Důležitým závěrem tohoto příspěvku je rovněž poznatek, že výzkum kladoucí si za cíl přemostění těchto oddělených světů se již dnes neobejde bez nasazení interdisciplinárních přístupů, které rovněž umožňují

využívat poznatky a metody sociálních věd. Efektivním metodologickým ukotvením je v této souvislosti zejm. anglosasky pojmovaná aplikovaná informatická disciplína IS, která v posledních cca dvou až třech desetiletích absorbovala mnoho poznatků psychologie, sociologie a sociokulturní antropologie.

Tento metodologicky orientovaný příspěvek tak odpověděl na otázku, jakými cestami se lze k výše naznačenému cíli ubírat, a to zejm. představením vědecké metodologie, která byla (a je) využita ke zkoumání této problematiky na příkladu konkrétního výzkumného projektu. Příspěvek mj. také stručně představil inovativní metodu spojení návrhového a akčního výzkumu, jejíž nasazení není v existujících vědeckých pracích (ani na mezinárodní úrovni) zdaleka typické a svým způsobem předznamenává budoucí sbližování německé návrhové tradice výzkumu (BISE) s anglosaskou behaviorální tradicí (IS). Návrh této metody je jedním z vedlejších vědeckých přínosů mého výzkumu. Cílem tohoto příspěvku naopak nebyla prezentace konkrétních získaných dat a navržených artefaktů – tyto výstupy byly či budou obsahem jiných příspěvků.

Literatura

- ACM/AIS/IEEE-CS, 2005. *Computing Curricula 2005: The Overview Report*, Available at: http://www.acm.org/education/curric_vols/CC2005-March06Final.pdf.
- Avgerou, C., 2000. Information systems: what sort of science is it? *Omega*, 28(5), s. 567–579.
- Avison, D. & Elliot, S., 2006. Scoping the discipline of information systems. In *Information Systems: The State of the Field*. Chichester: Wiley, s. 3–18.
- Baskerville, R. & Myers, M., 2002. Information systems as a reference discipline. *MIS Quarterly*, 26(1), s. 1–14.
- Bate, S., 1997. Whatever Happened to Organizational Anthropology? A Review of the Field of Organizational Ethnography and Anthropological Studies. *Human Relations*, 50(9), s. 1147–1175.
- Bělohávek, F., 1996. *Organizační chování*, Olomouc: Rubico.
- Beynon-Davies, P., 2007. Informatics and the Inca. *International Journal of Information Management*, 27(5), s. 306–318.
- Buhl, H.U. et al., 2012. Business and information systems engineering: A complementary approach to information systems – What we can learn from the past and may conclude from present reflection on the future. *Journal of the Association of Information Systems*, 13(4), s. 236–253.
- Buhl, H.U. & Lehnert, M., 2012. Information systems and business & information systems engineering: Status quo and outlook. In *15th International Conference on Business Information Systems, BIS 2012; Vilnius*. LNBIP, Springer, s. 1–10.
- Buchalceková, A., 2009. *Metodiky budování informačních systémů*, Praha: Oeconomica.
- Carr, N.G., 2003. IT doesn't matter. *Harvard business review*, 81(5), s. 41–9, 128.
- Cater-Steel, A. & Toleman, M., 2008. The impact of national culture on software engineering practices. *International Journal of Technology Policy and Management*, 8(1), s. 76–90.
- Czarniawska, B., 2012. Organization theory meets anthropology: a story of an encounter. *Journal of Business Anthropology*, 1(1), s. 118–140. Dostupné z: <http://ej.lib.cbs.dk/index.php/jba/article/view/3549>.
- Davison, R.M. & Martinsons, M.G., 2011. Methodological practice and policy for organisationally and socially relevant IS research: an inclusive–exclusive perspective. *Journal of Information Technology*, 26(4), s. 288–293.
- Doležel, M., 2013a. An initial proposal for a test governance framework in business organizations. In *IDIMT-2013, Information Technology, Human Values, Innovation and Economy*. Linz: Universitätsverlag Rudolf Trauner, s. 233–240.
- Doležel, M., 2013b. Kudy k dinosaurům? O behaviorálním paradigmatu v podnikové informatice, paralelách k antropologii organizací a svazujících hranicích oborů. *AntropoWebzin*, 9(4), s. 135–144.
- Geertz, C., 2000. *Interpretace kultur: vybrané eseje*, Praha: Sociologické nakladatelství.

- Hansen, B., Rose, J. & Tjørnehøj, G., 2004. Prescription, description, reflection: the shape of the software process improvement field. *International Journal of Information Management*, 24(6), s. 457–472.
- Hejnal, O., 2013. „Já si je najdu, ty vole. Dyt' máme furt stejný místa“: Místa bezdomovců ve veřejném prostoru. *Lidé města*, 15(3), s. 419–441.
- Von Hellens, L.A., 1997. Information systems quality versus software quality. A discussion from a managerial, an organisational and an engineering viewpoint. *Information and Software Technology*, 39(12), s. 801–808.
- Hevner, A. & Chatterjee, S., 2010. *Design Research in Information Systems: Theory and Practice*, New York: Springer.
- Hrešanová, E., 2008. *Porod z perspektivy sociálních věd: etnografie dvou českých porodnic se zaměřením na jejich (organizační) kultury*, Doktorská disertační práce FSS MUNI. Dostupné z: <http://is.muni.cz/th/133537/>.
- Chytil, M., 2010. Svět algoritmů a svět firem. In *Hovory s informatiky 2010*. Praha: Ústav informatiky AV ČR, s. 43–55.
- Jordan, A.T., 1994. Organizational culture: The anthropological approach. *Practicing anthropology in corporate America: consulting on organizational culture*, 14(1), s. 3–16.
- Kunda, G., 2009. *Engineering Culture: Control and Commitment in a High-Tech Corporation*, Temple University Press.
- Lukášová, R., 2010. *Organizační kultura a její změna*, Praha: Grada.
- Marrewijk, A. Van, Veenswijk, M. & Clegg, S., 2010. Organizing reflexivity in designed change: the ethnoventionist approach. *Journal of Organizational Change Management*, 23(3), s. 212–229.
- Mintzberg, H., 1968. *The manager at work – determining his activities, roles and programs by structured observation*, Cambridge, MA, USA: Doktorská disertační práce MIT. Dostupné z: <http://dspace.mit.edu/handle/1721.1/14146>.
- Moeran, B., 2009. From participant observation to observant participation. In *Organizational Ethnography: Studying the Complexities of Everyday Life*. London: SAGE Publications Ltd, s. 139–155.
- Molnár, Z. et al., 2012. *Pokročilé metody vědecké práce*, Praha: Profes Consulting.
- Morrison, J. & George, J., 1995. Exploring the software engineering component in MIS research. *Communications of the ACM*, 38(7), s. 80–91.
- Myers, M.D., 2009. *Qualitative Research in Business & Management*, London: Sage.
- Myers, M.D. & Tan, F.B., 2002. Beyond models of national culture in information systems research. *Journal of Global Information Management*, 10(1), s. 24–32.
- Nakonečný, M., 2005. *Sociální psychologie organizace*, Praha: Grada.
- Nový, I., Surynek, A. & kol., A., 2006. *Sociologie pro ekonomy a manažery 2. vydání*, Praha: Grada.
- Orlikowski, W. & Barley, S., 2001. Technology and institutions: what can research on information technology and research on organizations learn from each other? *MIS quarterly*, 25(2), s. 145–165.
- Osman, R. et al., 2011. K čemu doktorandi? Pokračování diskuze XXII. Sjezdu české geografické společnosti v Ostravě. *Informace ČGS*, 30(2), s. 18–24.
- Österle, H. et al., 2011. Memorandum on design-oriented information systems research. *European Journal of Information Systems*, 20(1), s. 7–10.
- Pavlica, K., 2000. *Sociální výzkum, podnik a management: průvodce manažera v oblasti výzkumu hospodářských organizací*, Praha: Ekopress.
- Pentland, B.T. & Feldman, M.S., 2008. Designing routines: On the folly of designing artifacts, while hoping for patterns of action. *Information and Organization*, 18(4), s. 235–250.
- Sein, M. et al., 2011. Action design research. *MIS quarterly*, 35(1), s. 37–56.

- Schein, E.H., 1996. Culture: The missing concept in organization studies. *Administrative Science Quarterly*, 41(2), s. 229–240.
- Schein, E.H., 2010. *Organizational Culture and Leadership* 4th ed., San Francisco, CA: Jossey-Bass.
- Sommerville, I., 2007. *Software Engineering* 8th ed., Harlow, England: Addison-Wesley.
- Svatá, V., 2012. *Audit informačního systému* 2. vydání., Příbram: Professional Publishing.
- Vlasák, R., 2009. Obor informační a inženýrský – několik poznámek z terminologické historie. *Ikaros*, 13(9). Available at: <http://ikaros.cz/node/5687>.
- Vodáček, L. & Rosický, A., 1997. *Informační management. Pojetí, poslání a aplikace*, Praha: Management Press.
- Voříšek, J. & Pour, J., 2012. *Management podnikové informatiky*, Prague: Professional Publishing.
- Walsham, G., 2006. Doing interpretive research. *European Journal of Information Systems*, 15(3), s. 320–330.
- Walsham, G., 1993. *Interpreting Information Systems in Organizations*, Chichester: John Wiley & Sons.
- Walsham, G., 1995. The emergence of interpretivism in IS research. *Information systems research*, 6(4), s. 376–394.
- Watson, T.J., 2011. Ethnography, Reality, and Truth: The Vital Need for Studies of “How Things Work” in Organizations and Management. *Journal of Management Studies*, 48(1), s. 202–217.
- Wilde, T. & Hess, T., 2007. Forschungsmethoden der Wirtschaftsinformatik: Eine empirische Untersuchung. *Wirtschaftsinformatik*, 49(4), s. 280–287.
- Wilson, G. & Aranda, J., 2011. Empirical Software Engineering. *American Scientist*, 99(6), s. 466–473. Available at: <http://www.americanscientist.org/issues/pub/empirical-software-engineering>.

JEL Classification: M15, M14, O15.

Summary

Behavioral and social factors in organizational IT processes: examination of their influence on software quality management and information systems deliveries

The aim of this methodological paper is a circumscription of a research project, focused on an intertwined concept of software quality management (with an emphasis on testing) and management of information systems development outsourcing. The research project is a design-centered one, with the main goal of designing a process framework for the purpose of business organizations. Another goal is to explore social and behavioral factors, in the same context. The paper briefly discusses the perspectives of Business Informatics and Software Engineering. The latter is the “traditional home” of software quality management concept. At the same time, the paper deals with the methodological dissimilarities between the fields of Business Informatics and Information Systems. The main thesis of the paper is to demonstrate that software should be perceived as a substantial part of information systems in organizations, however not the only part. The paper also clarifies the grounds which have guided the author, resulting in the conclusion of integrating the social and behavioral research part with the design one. Following this conclusion, the paper finally disseminates a new innovative research method. Such method, called “action design research”, is an ideal option for a similar category of research problems.

Key words: information systems development, software quality management, vendor management, outsourcing, action design research.

Řízení Informační bezpečnosti na základě standardního chování

Martin Dvořák

xdvom31@vse.cz

Doktorand oboru Aplikovaná Informatika

Školitel: doc. Ing. Zora Říhová, CSc. (rihova@vse.cz)

Abstrakt: Článek je zaměřen na problematiku řízení informační bezpečnosti organizace na základě vyhodnocování chování uživatelů, což je jeden z příkladů aplikace multiagentních systémů v praxi. Průzkumy (viz např. společnost Ernst&Young) dlouhodobě svědčí o vzrůstajícím zájmu organizací o alternativní systémy zajištění informační bezpečnosti. Důvodem jsou moderní trendy v IT a v informační bezpečnosti. V oblasti IT hovoříme o vzrůstající mobilitě a přibývajícím množství zařízení, které zpracovávají data organizace, např. tzv. „Internet of things“. V oblasti bezpečnosti IT se již několik let po sobě mezi prioritami organizací objevuje ochrana před ztrátou dat či jejím neoprávněným užitím a modifikací. To potvrzuje obecný trend společností o zajištění informační bezpečnosti nikoliv na základě vymezeného bezpečnostního perimetru ale na základě ochrany samotných dat.

Cílem této práce je ověřit, zda je možné za současných technických podmínek definovat model fungování tohoto alternativního systému, který je zaměřen na vyhodnocování uživatelského chování. Tento systém se zaměřuje primárně na ochranu dat a nikoliv na bezpečnostní perimetr a vyhodnocuje chování uživatelů v reálném čase. Dalším cílem je definování předpokladů pro jeho efektivní fungování. Vědecké metody použité v práci jsou primárně analýza a následně syntéza, dále statistické metody.

Klíčová slova: bezpečnost IT, řízení bezpečnosti, bezpečnostní incident, multiagentní systémy, standardní chování.

Úvod

Vlivem globalizace, moderních trendů v oblasti IT (virtualizace, cloud computing, mobilní aplikace, BYOD - Bring Your Own Device atd.) narůstajícím množstvím outsourcingových kontraktů a následným propojováním systémů mezi jednotlivými partnery se z oblasti řízení bezpečnosti informací stává stále komplexnější problém, jehož řešení nabývá na významu. Závěry z každoročně prováděných průzkumů ukazují, že rozdílnost mezi narůstajícími hrozbami a odpověďmi organizací na ně se zvyšují v exponenciální míře viz např. [3]. Bezpečnost IT z pohledu standardů definují standardy ČSN ISO/IEC řady 27000, které přijalo mnoho organizací. I přes toto zajištění bezpečnosti však dochází k únikům dat či bezpečnostním incidentům zaviněným lidským chováním. Lidský faktor (úmyslně či neúmyslně) je významným činitelem stojícím na počátku bezpečnostního incidentu, proto organizace v praxi začínají zvažovat alternativní metody řízení bezpečnosti s cílem zamezit úniku dat. Jednou z takových metod je i řízení informační bezpečnosti na základě standardního chování.

V této práci jsou použity statistické metody pro definování závěrů z průzkumů a dále autor používá analýzu a syntézu.

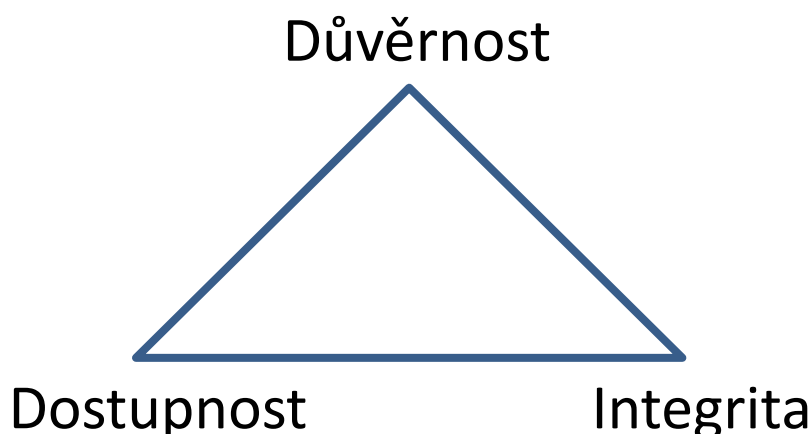
Základní předpoklady a definice základních pojmů

Autor předpokládá základní znalost pojmů v oblasti bezpečnosti IT, které jsou definovány mezinárodními normami ČSN ISO/IEC 27001 [1]. V této práci jsou zmiňovány zejména tyto pojmy:

- **Standardní chování** můžeme definovat jako takové chování, které je nutné k vykonávání určité svěřené pozice v organizaci a plnění úkolů z ní plynoucích. Takovéto chování je v souladu se všemi interními předpisy a směrnici organizace, zákony a etickými normami.
- **Bezpečnost IT** lze definovat pomocí normy ČSN ISO/IEC 27001 [1], která tento pojem popisuje takto: „Bezpečnost informací je zachování důvěrnosti, integrity a dostupnosti informací a

dalších vlastností jako např.: autentičnost, odpovědnost, nepopiratelnost a spolehlivost¹. Někdy bývá označována první část této definice jako tzv.: CIA1 trojúhelník bezpečnosti – viz obrázek 1.

- Obecně je za **bezpečnostní incident** považována událost, která zahrnuje porušení bezpečnosti (dostupnosti, integrity či důvěrnosti) informací. To může být např.: porušení směrnic společnosti, norem, zákonů nebo nepostupování obsluhy výpočetního zařízení v souladu s návodem k němu². Z úhlu pohledu této práce (vzhledem k jejímu zaměření), budeme za bezpečnostní incident považovat primárně jakékoliv chování mimo standardní chování uživatele.



Obrázek 7: Definice informační bezpečnosti

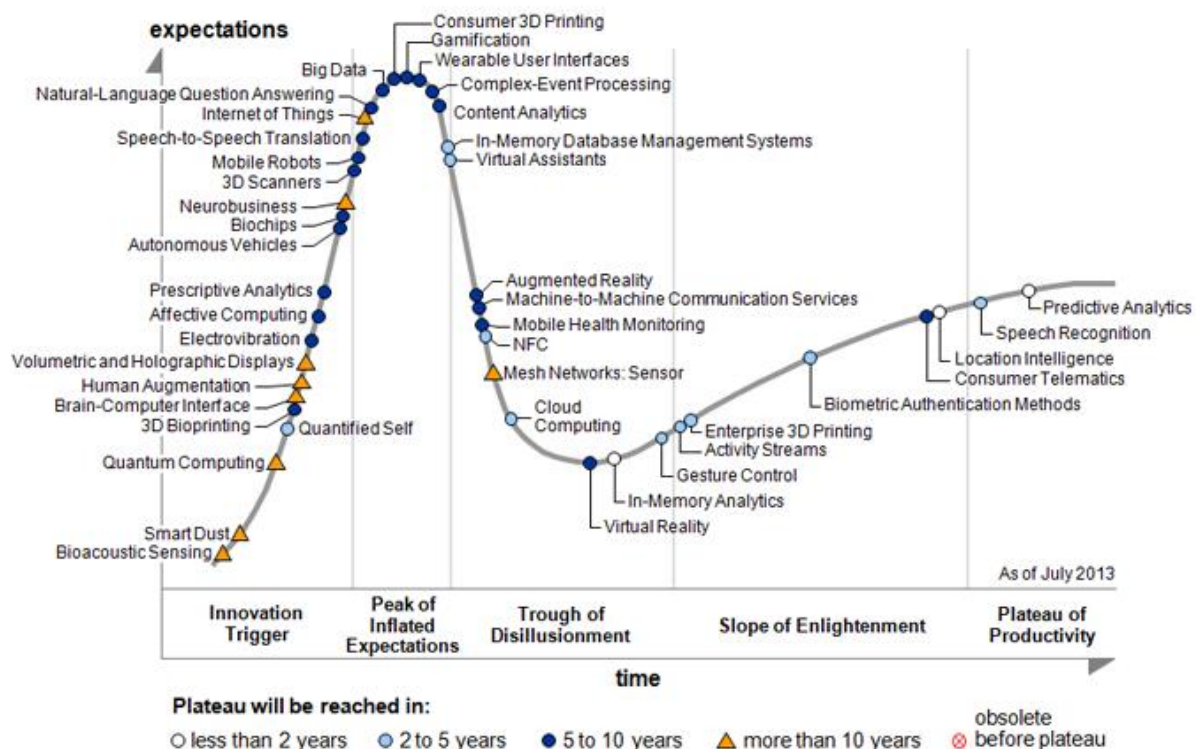
Zdroj: Autor dle ČSN ISO/IEC 27001 [1]

Trendy v oblasti informační bezpečnosti

Současný přístup k zajištění informační bezpečnosti spočívá v provedení několika kroků a jejich pravidelnému opakování (tzv. spirála zlepšování). Organizace nejprve provede rizikovou analýzu pro identifikaci všech hrozeb a následně definuje a realizuje opatření, která mají minimalizovat úroveň rizika na přijatelnou úroveň. V praxi velmi rozšířený přístup k zajištění informační bezpečnosti spočívá v oddělení firemní infrastruktury a dat od té veřejné „internetu“, což znamená „schovat“ vše firemní důvěrné za firewall. Tento přístup díky moderním trendům v oblasti IT přestává plně uspokojovat všechny požadavky organizace (především na mobilitu), proto je nutné vyvinout nový způsob zajištění bezpečnosti, který bude tyto trendy reflektovat. Moderní trendy v oblasti IT zobrazuje obrázek č. 2. Z tohoto obrázku mají největší vliv na bezpečnost trendy Internet of Things, Machine to machine komunikace, Big Data, Autonomní automobily a Virtuální asistenti. Všechny tyto trendy směřují k faktu, že stále více zařízení bude mít přístup k datům organizace, budou je zpracovávat a ve většině případů budou mobilní, což znamená, že mohou snadno opustit vymezené hranice systému řízení bezpečnosti organizace. Dalšími fakty, které přispívají ke snaze řešit bezpečnost alternativními způsoby jsou: objevující se škodlivé programy, které umí zneužít mobilních zařízení k přenosu přístupových údajů i bez připojení k internetu, např. jeden z nich vyvinuli vědci z Fraunhoferova institutu v Německu – viz: [7]. Druhým faktem je, že bezpečnost je ve většině organizací velmi dobře zajištěna na úrovni fyzických a technických opatření, avšak slabým místem se stává člověk a jeho selhání (úmyslné či neúmyslné) – viz např.: [2]. Toto také potvrzují trendy na obrázku 3, kdy hrozba úniku dat se za posledních několik let vždy dostala do popředí výsledků průzkumů priorit v oblasti bezpečnosti. Třetím faktem je neustále rostoucí počet připojených zařízení k internetu: v roce 2012 byl počet připojených zařízení 8,7 miliardy, v roce 2013 tento počet překročil hranici 10 miliard a pro rok 2020 se dle studie společnosti Cisco [5] očekává 50 miliard připojených zařízení.

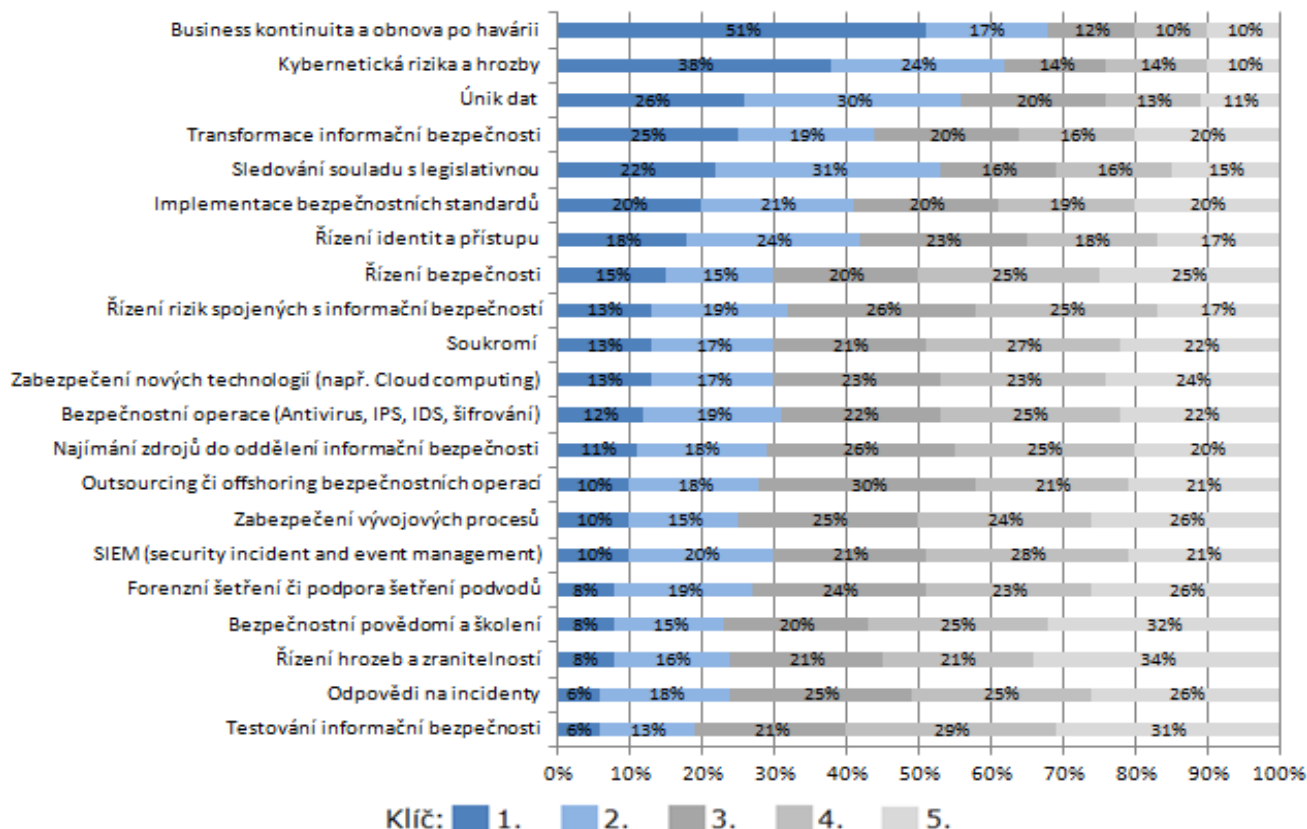
¹ CIA z anglických slov Confidentiality, Integrity a Availability.

² Bezpečnostní incident může nastat aniž by byl přímo škodlivý nebo by byla přímo patrná škoda v momentě jeho nastání. Vše se může projevit později.



Obrázek 8: Gartner Hype Cycle trendů v oblasti IT pro rok 2013

Zdroj: [6]



Obrázek 9: Nejvýznamnější priority v oblasti informační bezpečnosti

Zdroj: Ernst&Young, 2013, [4]

Řízení informační bezpečnosti na základě standardního chování

Výše uvedené trendy a hrozby v oblasti IT a informační bezpečnosti poskytují příležitost k nalézání alternativních způsobů zajištění informační bezpečnosti. Tyto alternativní způsoby se na rozdíl od toho klasického, jehož cílem je vymezení a zabezpečení bezpečnostního perimetru, snaží reflektovat mobilitu – tzn. cílem je ochrana dat a nikoliv perimetru. Jedním z takových přístupů je i řízení informační bezpečnosti na základě standardního chování.

Základní způsob fungování tohoto přístupu k bezpečnosti IT je počáteční definice standardního chování, monitoring chování všech uzlů v systému, neustálé vyhodnocování tohoto chování a následné přijetí akce (pokračovat či přerušit akci), tak aby nemohl žádný uzel systému provést škodu v podobě zcizení dat, jejich smazání či neoprávněné manipulace s daty. Jedná se o aplikaci multiagentních přístupů v praxi.

Jak naznačuje odstavec výše, přistupujeme ke všem prvkům v systému jako k uzlům (agentům). Předpokládáme tedy, že každý prvek (uživatel, klient, uzel...) v systému má díky svému účelu³ určitý model svého chování. Pokud se tento model podaří definovat, lze jej porovnávat s pozorováním skutečného chování. Na základě tohoto porovnání bude systém schopen poznat, zda se jedná o:

- známý povolený stav,
- známý mimořádný („nepovolený“) stav,
- neznámý mimořádný stav (tyto stavy by měly jít následně vyhodnotit a zahrnout do definice modelu⁴).

Současné poznání v tomto oboru již pracuje podobně – jsou to systémy IPS, IDS⁵ a SIEM⁶ sloužící k prevenci a detekci průniků do sítě. Pokud se však chceme bavit o zajištění informační bezpečnosti ve smyslu její plné definice, nejsou tyto systémy schopny zajistit integritu a plnohodnotně nezabezpečují důvěrnost dat. Pokud bychom měli demonstrovat posun, který nabízí zajištění informační bezpečnosti pomocí standardního chování, lze použít ISO OSI model standardizovaný jako ISO 7498 [8]. Tento model počítá s celkem 7 vrstvami s cílem propojování otevřených systémů. Zatímco současné systémy pracují primárně do vrstvy 4, maximálně 5, systém vyhodnocující chování musí pracovat v úrovni 6 – 7, čili v prezenční až aplikační vrstvě. Důvodem, proč současné systémy nemohou pracovat výše je šifrování (např. SSL) již na úrovni aplikací. V důsledku tohoto mohou současné systémy posoudit oprávněnost komunikace klient – server, ale již nejsou schopni posoudit, zda data uvnitř této komunikace jsou poskytována oprávněně či nikoliv.

Definice modelu řízení informační bezpečnosti

Model řízení informační bezpečnosti na základě standardního chování je v podstatě praktickou aplikací multiagentního přístupu k systému (modelu). Autor hodlá k definici tohoto modelu použít „standardní“ nástroje. Pro komunikaci agentů navzájem čili jazyk ACL (Agent Communication Language) bude sloužit jazyk KQML (Knowledge Query and Manipulation Language) viz [9]. Tento jazyk primárně slouží k:

- identifikaci vhodných agentů pro spolupráci
- navazování spojení mezi agenty
- výměně informací a dat mezi agenty

Pro vlastní obsah sdělení (vyměňované informace a data) autor navrhuje použít standardní XML jazyk.

Abychom mohli model navrhnout a později s ním pracovat a ověřovat jej v praxi je nutné definovat architekturu modelu. Autor hodlá využít v této oblasti nejznámější referenční architekturu agentů: BDI (Belief-Desire-intention) viz [MARIK02]. Architektura pracuje se třemi základními atributy agenta:

- Beliefs (domněnky) – informace o okolním světě agenta,
- Desires (přání) – souhrn motivací, cílů a cílových stavů jeho chování

³ U lidí nehovoříme pouze o účelu, ale také o zájmu či motivaci.

⁴ Tzv. proces učení.

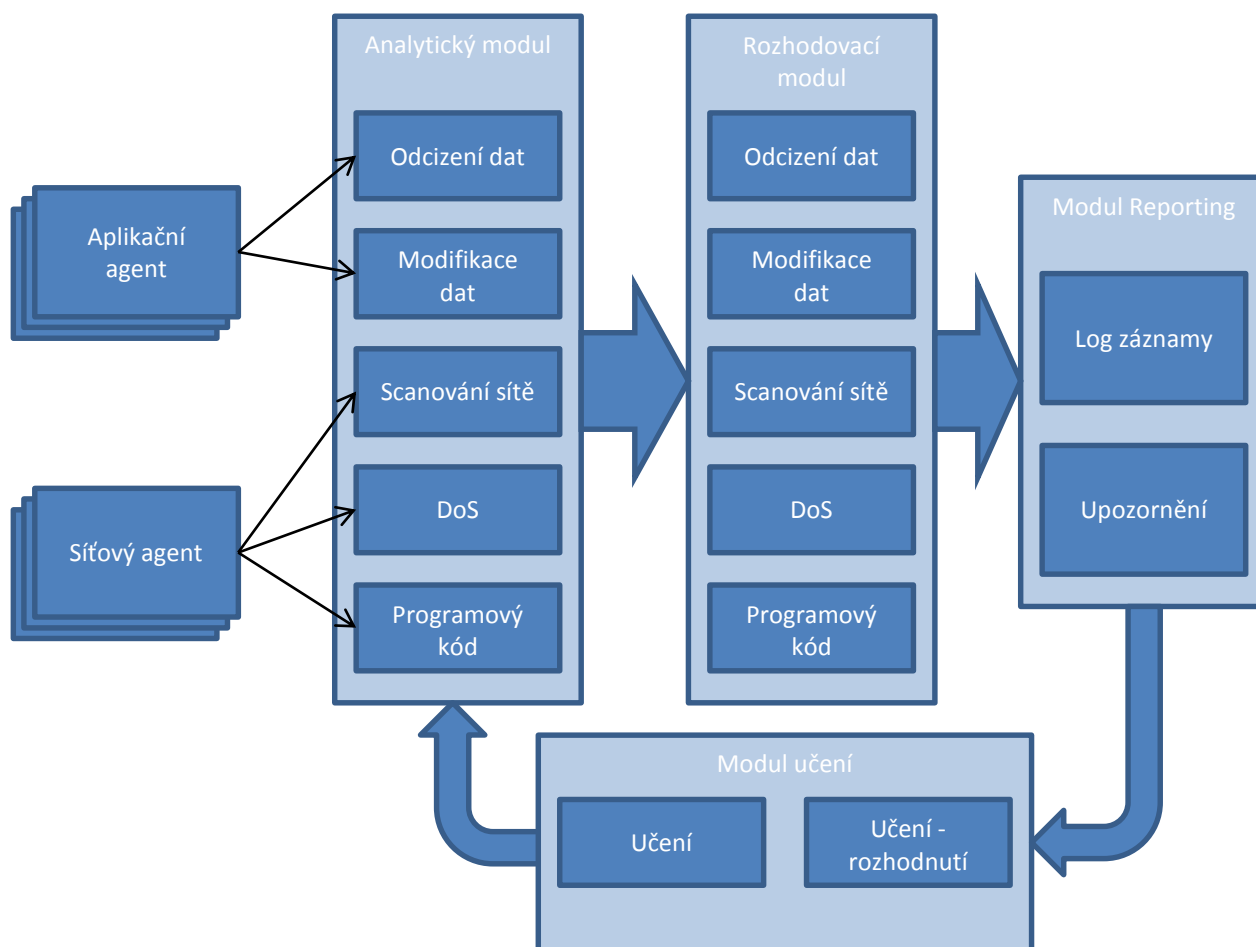
⁵ Intrusion Prevention System a Intrusion Detection System.

⁶ Security Information and Event Management.

- Intentions (záměry) – rozhodnutí agenta, jak ovlivní vývoj v blízké budoucnosti, aby bylo dosaženo nějakého přání. Záměr tedy je určitý lokální cíl agenta.

K zajištění informační bezpečnosti systémem, který umí vyhodnocovat standardního chování je nutné nastavit počáteční definici chování a kategorizaci úrovně ochrany dat. Systém si následně bude definici chování stále díky funkci učení zajišťovat aktuální. Problematika modelování chování není předmětem tohoto článku, přesto však lze v krátkosti shrnout několik základních přístupů:

- Ex Post přístup spočívá v definici standardního chování na základě minulosti, tj. událostí, které se již staly. Např. lze k tomuto účelu použít data-mining přístupy nad různými datovými stopami uživatele – např. log soubory z různých aplikací organizace, odchytáváním datových paketů na síti⁷ nebo monitoringem současného stavu. Zde je riziko zanesení nestandardních přístupů do definice standardního chování.
- Ex Ante přístup spočívá v definici standardního chování do budoucna např. z popisu procesů a rolí v nich, na základě zkušeností konzultantů, kteří definici vytvářejí apod.



Obrázek 10: Model systému řízení informační bezpečnosti na základě vyhodnocování standardního chování

Zdroj: Autor

K tomu, aby systém řízení bezpečnosti IT pomocí standardního chování uživatelů viz obrázek 4 byl účinný, musí organizace, kde je nasazen, splňovat určité předpoklady. Níže definované předpoklady vychází z analýzy fungování současných systémů, definovaného modelu a jeho funkcí, dnešních trendů v oblasti IT a v uspořádání organizací. Je však možné, že během testování přibudou další předpoklady. Na základě tohoto lze určit následující předpoklady:

- Jasná definice rolí zaměstnanců „uživatelů“ pro stanovení modelu chování uživatelů
- Jasná definice procesů v organizaci pro stanovení modelu chování uživatelů

⁷ Při analýze auditních záznamů (log souborů) se zpravidla nedozvíme všechna konkrétní transakční data, ale lze z těchto dat získat informace o událostech, které uživatel vyvolal.

- Dostatečná kvalifikace jedince vykonávající daný proces
- Velikost organizace
- Podpůrné nástroje pro vykonávání procesů
- Velikost datové základny firmy
- Dostupnost a strukturovanost veřejných dat k analýze

V modelu na obrázku 4 vystupují agenti v následujících rolích:

- Aplikační agent – Reaktivní
- Síťový agent – Reaktivní
- Agenti pro analýzu chování typu:
 - Odcizení dat – Proaktivní agent
 - Modifikace dat – Proaktivní agent
 - Scanování sítě – Proaktivní agent
 - DoS (Denial of Service) – Proaktivní agent
 - Neoprávněný programový kód – Proaktivní agent
- Agenti pro rozhodování:
 - Odcizení dat – Proaktivní cílově orientovaný agent
 - Modifikace dat – Proaktivní cílově orientovaný agent
 - Scanování sítě – Proaktivní cílově orientovaný agent
 - DoS (Denial of Service) – Proaktivní cílově orientovaný agent
 - Neoprávněný programový kód – Proaktivní cílově orientovaný agent
- Reportovací agenti:
 - Log záznamy – Reaktivní agent
 - Upozornění – Reaktivní agent
- Modul Učení:
 - Učení – Rozhodnutí – Reaktivní Agent
 - Učení – Reaktivní Agent

Závěr

Cílem článku byla definice modelu pro zajištění informační bezpečnosti pomocí vyhodnocování standardního chování. Důvodem tohoto alternativního přístupu k zajištění informační bezpečnosti jsou trendy v oblasti IT (vzrůstající mobilita a počet zařízení zpracovávajících data organizace) a trendy v oblasti informační bezpečnosti (primárně snaha o ochranu proti úniku dat) a zvažování.

Model výše zmíněné trendy zohledňuje a posouvá úroveň řízení informační bezpečnosti až do aplikační úrovně modelu ISO OSI. Díky tomuto je dosažena ochrana dat proti odcizení či neoprávněné manipulaci a to i v šifrovaném prostředí, které znamená přenos paketů v zašifrované podobě.

Autor chce dále prověřovat fungování modelu v praxi, buď jeho praktickým nasazením nebo na základě simulací nad daty z reálného provozu.

Praktické využití se týká především organizací, kde ztráta dostupnosti či integrity dat může mít fatální následky pro její další fungování: banky, státní instituce, vědecké a zdravotní instituce, telefonní operátoři, pojišťovny apod.

Literatura

[1] ČSN ISO/IEC 27001 Informační technologie – bezpečnostní techniky – systémy managementu bezpečnosti informací – Požadavky.

[2] DVOŘÁK, Martin, ŘÍHOVÁ, Zora. *Problems of ISO 27001 Matrix certification*. Prague 25.05.2011 – 26.05.2011. In: *Information Security Summit*. Praha: TATE International, 2011, s. 41–58. ISBN 978-80-86813-22-6.

[3] ERNST & YOUNG. Global Information Security Survey 2012: *Fighting to close the gap*. [cit. 11. Ledna 2014] Dostupný online: <<http://www.ey.com/Publication/vwLUAssets/>

Fighting_to_close_the_gap:_2012_Global_Information_Security_Survey/\$FILE/2012_Global_Information_Security_Survey___Fighting_to_close_the_gap.pdf>

[4] ERNST & YOUNG. Global Information Security Survey 2013: *Under Cyber Attack*. [cit. 11. Ledna 2014] Dostupný online: <[http://www.ey.com/Publication/vwLUAssets/EY_-_2013_Global_Information_Security_Survey/\\$FILE/EY-GISS-Under-cyber-attack.pdf](http://www.ey.com/Publication/vwLUAssets/EY_-_2013_Global_Information_Security_Survey/$FILE/EY-GISS-Under-cyber-attack.pdf)>

[5] EVANS, Dave. *The Internet of Things: How the next evolution of the Internet is changing everything*. [cit. 12. Ledna 2014] Dostupný online: <http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf>

[6] GARTNER. *Gartner's 2013 Hype Cycle for Emerging Technologies Maps Out Evolving Relationship Between Humans and Machines*. [cit. 12. Ledna 2014] Dostupný online: <<http://www.gartner.com/newsroom/id/2575515>>

[7] GOODIN, Dan. *Scientist-developed malware prototype covertly jumps air gaps using inaudible sound*. [cit. 12. Ledna 2014] Dostupný online: <<http://arstechnica.com/security/2013/12/scientist-developed-malware-covertly-jumps-air-gaps-using-inaudible-sound/>>

[8] ISO/IEC 7498-1:1994. *Information Technology – Open Systems Interconnection – Basic Reference Model: The Basic Model*.

[9] MAŘÍK, Vladimír. ŠTĚPÁNKOVÁ, Olga. LAŽAŇSKÝ, Jiří. A kol. *Umělá inteligence* (3). 1. Vyd. Praha: Academia 2001. 328 s. ISBN: 80-200-0472-6

JEL klasifikace: L86, C88, L29.

Summary

Information security management based on standard behavior

This paper is focused on information security management of organization based on evaluation of standard user behavior, which is one of the examples of the multiagent system practical application. This alternative approach to information security management is based on data protection rather than security perimeter definition and its following protection. The defined model of such system is able to evaluate user behavior in real time and is focused on data protection. Given these facts the main requirements of business organizations are reflected: mobility and data leakage protection. These two main requirements are resulting from main trends in IT field (internet of things and mobile applications) and information security (data leakage protection).

Author would like to further assess and test the model in praxis or use real data simulation. As a result of this the presumptions for its effective operation will be further refined. Additional next step is to focus on definition of standard behavior and its simulation and evaluation.

Key words: IT security, security management, security incident, multiagent systems, standard behavior.

Přínosy agilních metodik řízení ICT projektů

Jan Juříček

jan.juricek@vse.cz

Doktorand oboru aplikovaná informatika

Školitel: doc. Ing. Vlasta Svatá, (svata@vse.cz)

Abstrakt: Článek shrnuje historický vývoj metodik používaných při řízení projektů. Shrnuje současné zavedené rámce procesních metodik řízení projektu PMI a Prince2 a konfrontuje je s moderními agilními přístupy. Agilní přístupy a agilní způsob řízení projektů je v článku vymezeno historicky i obsahově. Historicky v rámci reakcí na překonané vodopádové modely řízení a obsahově v podobě agilních principů. Agilní přístupy shrnují základní principy, kterými by se měl úspěšný projekt řídit. Konfrontace s metodikami Prince2 a PMI spočívá v identifikaci potenciálních problémů spojených při využívání rigorózních metod a definici jejich příčin v rámci modelů procesních metodik. Při porovnání s jednotlivými fragmenty procesních metodik, kterými jsou například standardizované procesy, principy, směricové postupy a domény, akcentuje článek agilní principy řízení projektů a tím i přínosy agilních technik při řízení ICT projektů. Dále se práce věnuje jednotlivým druhům agilních metodik, kterými jsou například extrémní programování (XP), SCRUM nebo Lean Development.

Klíčová slova: projektové řízení, metodiky, agilní řízení.

Úvod do historie projektového řízení

Projektové řízení (PŘ) se jakožto specializovaná disciplína začalo vyvíjet až v druhé polovině 20. století (Kwak, 2005). Do té doby bylo spojováno pouze s civilním inženýrstvím a stavitelstvím při definici, v průběhu komplexních inženýrských staveb včetně technických popisů. Až ve druhé polovině se disciplína PŘ začala rozrůstat i o projekty¹ v oblasti strojírenství, výroby a zbrojního průmyslu.

Dvěma hlavními představiteli v tomto období jsou Henry Gantt a Henry Fayol. Gantt je představitel detailního plánování a kontrolních technik – již zažitý Gantt diagram nese právě jeho jméno. Fayol pak zformulování 5 klíčových aktivit projektového managementu. Oba pak do metodik prosazovali použití „Work Breakdown Structure“ tzn. rozpadu projektu jako detailního artefaktu plánování výstupů a zdrojů na projektu (Cleland, Gareys, 2006). V těchto raných letech PŘ začínalo být jasné, že pro kvalitní naplánování a tedy i pro iniciaci úspěšného projektu je potřeba definovat a prioritizovat požadavky a specifikovat rozsah projektu.

V 50. letech byly Ganttovy diagramy, modely CPM a PERT hlavními plánovacími a kontrolními technikami. CPM – Metoda kritické cesty je algoritmus pro naplánování projektových aktivit (Kelley, 1961). CPM udává takovou sekvenci aktivit, která určuje délku trvání projektu. Prakticky počítá nejdelší cestu mezi jednotlivými úkoly v projektu a zvýrazňuje tím kritické body – kritické úkoly v projektech. Zdržení na kritických úkolech znamenají prodloužení – zpoždění milníku nebo celého projektu. PERT („Project Evaluation and Review Technique“) vznikl jako statistický nástroj v rámci armády spojených států. Tato metoda detailně analyzuje veškeré úkoly (činnosti) v projektu a počítá se třemi scénáři v rámci doby trvání na každém úkolu – optimistickým, reálným a pesimistickým (Milosevic, 2003).

V šedesátých letech došlo k vytvoření dvou profesionálních asociací pro projektové řízení – IPMA (International Project Management Association) jakožto hlavní evropský proud (společně s Anglickým Prince2) a PMI (Project Management Institute) v USA. Obě skupiny razily trend vodopádových metodik v podobě pevně definovaných procesů v tzv. tradičním přístupu (Fernandez, 2009):

1. Inicie
2. Plánování a definice
3. Exekuce
4. Monitorování a kontrola
5. Dokončení

¹ Projekt je dočasné úsilí k vytvoření jedinečného produktu, služby nebo výsledku. Vyznačuje se definovaným začátkem a koncem, jedinečností a měřitelností (PMI, 2010)

V druhé polovině dvacátého století se paralelně rozvíjel v Japonsku i další trend, který zasel první semínko agilních přístupů v projektovém řízení. Jednalo se o zeštíhlení výroby v Japonské Toyotě, kde se praktikoval plně přístup „naš zákazník, náš pán“ a bylo nutné v maximální míře uspokojit požadavky – zapracované do produktů v co možná nejkratší době a pokud možno s minimálními náklady. Tato metodika se nazývá LEAN a akcentuje právě vytvořenou hodnotu pro koncového zákazníka a hledání dokonalosti skrze neustálý cyklus zlepšování (Boyle, 2011).

V devadesátých letech spatřila světlo světa metoda, která opět posunula hranice možností zkracování délky milníků a byla velmi vřele přijata nejen akademiky ale i odbornou veřejností a znamenala největší posun v projektovém řízení od CPM. Jednalo se o Teorii kritického řetězce autora Eliyaha Goldratta akcentující: redukci multitaskingu, celkovou dobu projektu na dostupnosti projektových zdrojů (nejen na kritické cestě), potlačuje Parkinsonův zákon² v projektovém řízení, potlačuje také studentský syndrom - Zdroje mají tendenci zahajovat projektovou činnost "na poslední chvíli" - a mají k tomu spoustu objektivních důvodů, které většinou ani sami nemohou příliš ovlivnit. Ovšem pravděpodobnost nedodržení termínu projektové činnosti se tím astronomicky zvyšuje. Naopak metoda akcentuje princip štafetového běžce (kritický řetěz nastavuje pomocí nástrojů řízení projektu situaci, kdy jednotlivé zdroje po zahájení práce na projektové činnosti "běží co nejrychleji" a jakmile činnost dokončí, předají ji okamžitě dále bez ohledu na termín) a také nutí manažera sledovat průběh činnosti nikoliv na základě milníků, ale na základě zbytkových časů v náraznících před milníky a tím prioritizovat činnosti a přidělovat krizové plány (Goldratt, 1997).

Přístupy a metodiky

Při bližším zkoumání procesních metodik Prince2, PMI a IPMA nalezneme několik vzájemných shod, které determinují správný průběh projektu. Jedná se primárně o dodržování klíčových procesů projektového řízení. Správa společnosti Standish Group z roku 2004, na základě šetření u 23 společností různých velikostí udává další dva závěry. Prvním je zjištění, že zkušenosti projektového manažera korelují s dodržáním rozpočtu a harmonogramu projektu. Druhým zjištěním je pak fakt, že větší projekty (1000 člověkodnů a více) jsou na tom z úspěšností v průměru hůře než projekty menší.

Klíčové procesy byly identifikovány na základě dotazníků, kdy škála důležitosti procesu byla definována od 1 do 5, kdy 5 znamenala nejkritičtější proces. Klíčovými procesy správného průběhu projektu jsou (Standish Group, 2004):

Tabulka 1: Klíčové procesy

Zdroj: Standish Group 2004

Pořadí důležitosti	Proces v projektovém řízení	Průměrná důležitost
1	Monitorování a kontrola	4,81
2	Řízení změn rozsahu	4,77
3	Podpora top managementu	4,77
4	Skladba projektového týmu	4,76
5	Odhadování zdrojů	4,73
6	Integrované řízení změn	4,73
7	Plné zapojení uživatelů	4,73
8	Deklarace rozsahu projektu	4,69
9	Kontrola plánu	4,69
10	Tvorba projektového plánu	4,69

PMI provedla ke konci roku 2013 obsáhlou studii mapující klíčové faktory úspěchu a selhání projektů. Zpráva zpracovala výsledky odpovědí více než 1000 seniorních projektových manažerů při řízení pro dané společnosti klíčových projektů. Zpráva udává 5 skupin klíčových faktorů pro úspěch projektu: personální nastavení, realistické plánování, podporu od nejvyšší úrovně, definování výhod a řízení změn (PMI, 2013). 10 klíčovými faktory podle uvedené studie pak jsou:

- Zapojení uživatelů,
- Podpora nejvyššího managementu a sponzora projektu,

² Parkinsonův projektový zákon: "Činnost trvá nejméně tak dlouho, jak dlouhý má přidělený časový interval."

- Jasný soupis požadavků,
- Správné plánování,
- Realistické očekávání,
- Kratší čas mezi projektovými milníky,
- Kompetentnost týmu,
- Přijmutí vlastnictví projektu,
- Jasná vize a cíle,
- Tvrdá a poctivá práce týmu a zaměření na cíl.

Metodika Prince2 se při klíčových faktorů úspěšnosti projektu opírá o definici SMART. Projekt musí být specifický, měřitelný, akceptovatelný, dosažitelný, realistický a časově definovaný – ukončitelný (OGC, 2011). Taktéž akcentuje (podobně jako PMI – realistické plánování) techniku produktového plánování PBS – podobné jako WBS, tedy rozpadnutí výsledných produktů na jednotlivé komponenty, jejich popis a kroky vedoucí k sestavení finálního produktu projektu.

Příčiny selhání IT projektů

Složitost a komplexnost projektů v oblasti informačních a komunikačních technologií, současné turbulentní prostředí globální ekonomiky i řada chyb, kterých se dopouštějí účastníci těchto projektů, má za následek krizi IT/IS projektů.

Studie Standish Group poukazuje na 10 hlavních faktorů, které způsobují selhání IT projektů:

- Nedostatek uživatelských vstupů,
- Nekompletní požadavky a specifikace,
- Změny v zadání, požadavcích, specifikacích,
- Nedostatek podpory vrcholného vedení,
- Technologické problémy,
- Nedostatek zdrojů,
- Nerealistické očekávání,
- Nejasné cíle,
- Nerealistické časové rámce dodání,
- Nové technologie, integrace technologie.

Studie Ernst&Young udává opakující se problémy, které způsobují selhání projektů: změna rozsahu projektu vlivem dvou faktorů – nepředpokládané (mnohdy externí) změny nebo neúplným plánováním; rozdílné představy o cílech a výstupech projektu; nedostatečná podpora managementu; optimistické očekávání a slabý rozpočet (Ernst&Young, 2011).

Procesní metodiky Prince2 a PMI prošly posledními 5 lety značným vylepšením, co se týče organizace práce – například Prince2 zavádí novou roli uvnitř organizačního diagramu postavenou na úroveň projektového manažera – „User Assurance“. Dále akcentuje více rozložení delších projektů na jednotlivé menší časové úseky „Managing Stage Boundaries“, ve kterých si ověřuje požadované stavy, ověřuje si přínosy dle Business Case, a navrhuje a doporučuje řídicímu výboru projektu další kroky.

Existují však postupy, se kterými jsou metodiky principiálně spoutány a které vzhledem k výše uvedeným studiím a výstupům nahrávají k neúspěchu. Procesní metodiky akcentují iniciační plánování, což vede k neefektivitě při změnách rozsahu projektu (rozpočet, termíny, kvalita). Změny pak musejí být velmi formálně dokumentovány a schvalovány. Zodpovědní sponzoři projektu se neradi adaptují na změnu. Bývají odtrženi od reality v dodávkách a jejich rozhodnutí jsou časově nákladná. Tato odtržení od reality jsou daná i principem „managing by exception“, jinými slovy typem řízení, kdy řídicí výbor projektu sleduje stav projektu a zasahuje do něj (a koriguje) pouze, pokud jsou překročeny nebo vychýleny dané metriky sledování. Nehledě na to, že sponzoři vidí občas své priority v jiných aktivitách a nepřikládají průběhu projektu prioritní význam.

Problémem nikoliv na straně procesních metodik, ale běžné praxe, může být neexistence vnímání rizikové složky rozpočtu. Jinými slovy, složka rozpočtu, která je alokována na předpokládaná rizika a jejich

mitigaci, není schvalována nebo je hluboko podceňována. Následkem pak je krizové řízení, nárůst byrokracie a dodatečných nákladů.

Pokud shrneme výše uvedená zjištění a akcentujeme zejména fakt, že mezi plánováním a dodáním existuje robustní časový rámec, ve kterém se může přihodit prakticky cokoliv, lze předpokládat, že vybrané principy agilních postupů budou mít na průběh projektů pozitivní účinek.

Agilní principy

Komunita okolo agilních přístupů ve vývoji se začala formovat v roce 2001, kdy se v Utahu sešlo několik uznávaných projektových manažerů a konzultantů a hledali spolu způsoby, jak řešit problémy tehdejších projektů. Výsledkem jejich schůzky bylo sepsání prohlášení, ve kterém vyzdvihli některé priority, podle kterých by se měl vývoj řídit. Šlo o „Manifesto for Agile Software Development“ (agile manifesto) a v něm říkají, že dávají přednost:

- individualitám a interakci nad procesy a nástroji,
- fungujícím SW před vyčerpávající dokumentací,
- spolupráci se zákazníkem před vyjednáváním o smluvních podmínkách,
- reakci na změnu nad plněním plánu za každou cenu (Beck, 2001).

Agilní přístupy jsou především o iteraci, kontinuální integraci a spolupráci. Agilností je v tomto kontextu vhodné rozumět jako rozdílným přístupem k řízení týmů a celých projektů. Agilní principy jsou:

- zahrnutí klíčového uživatele,
- zodpovědnost a síla v rozhodování týmu jdoucí z nižších pater managementu,
- požadavky se mění, časová osa zůstává,
- rychlé, malé, inkrementální skoky,
- nejdříve dodat jednu část kompletně, poté přejít na další – prioritizace požadavků a mitigace rizika,
- testovat dříve a úplně,
- spolupráce týmů, spolupráce sponzorů (Highsmith, 2002).

Spolupráce a zahrnutí uživatelů je klasický agilní imperativ. Uživatelé jsou nedílnou součástí dodavatelského týmu. V Agilních metodikách hraje velkou roli tzv. Vlastník produktu („Product Owner“), kdo je jednoznačně odpovědný za maximalizaci hodnoty výstupného produktu a celé týmové práce. Vedlejším efektem úzké spolupráce je pak velká vizibilita stavu celého projektu a méně byrokracie v reportingu.

Podle agilních metod je vlastnictví produktu (a zodpovědnost dodání produktu) utvořeno na úrovni řízení projektu, nikoliv na úrovni řídicího orgánu projektu – řídicího výboru (Watters, 2013). V reálné praxi se stává, že zásahy řídicího výboru ve věci měnících se požadavků redukuje motivace členů týmu dodávky. Agilním principem je skutečnost, že požadavky jsou prioritizovány a ohodnocovány na úrovni řešitelského týmu.

Práce s požadavky jsou obecně mantrou agilního pojetí řízení. V agilním vývoji se požadavky vyvíjejí, i když časová osa projektu zůstává neměnná. Tím vzniká extrémní zátěž na finanční plánování rozpočtu. Zde narážíme na úplně jinou filosofii sestavování rozpočtu. V případě WBS dokážeme rozpočet sestavit komponentně. Agilní metodika pracuje s celkovou sumou rozpočtu, vypočítanou na všechny členy týmu v rámci jedné iterace. V rámci jedné jediné iterace se provádí plánování, vyhodnocení rizika a samozřejmě prioritizace požadavků. Tato prioritizace vychází, dle světových kapacit v oboru, výhradně na obchodních benefitech konkrétního požadavku (Watters, 2013). Je tedy zřejmé, že agilní přístup můžeme aplikovat pouze na určité druhy projektů, kde budou pod kontrolou všechny externí dodavatelské složky, případně jasně nastavené řešitelské týmy. Na druhou stranu je též zřejmé, že čím více změn v požadavcích se dá očekávat, tím vhodnější bude nasazení agilnějších metodik do řízení projektu.

V rámci dodávek je agilním principem rychlost a možnost dílčího předvedení výstupu. Tím dokáže tým udržet soudržnost i vizibilitu na projektu a definitivně zkontrolovat funkčnost výstupu se zákazníkem. Délku iterace je vhodné nastavit na několik dní.

Výhody iterativního přístupu jsou dle Hightsmitha a Becka následující:

- redukce rizika (jasná vizibilita co je kdy na projektu hotového),
- zvyšující se hodnota – časté dodání výstupů,
- flexibilita a adaptabilita, na základě průběžného testování a změn funkčnosti dokáže řešitel pružně reagovat,
- řízení nákladů projektů, i v průběhu projektu již můžeme realizovat průběžné cíle a získávat předpokládanou hodnotu s parciálních výstupů (Hightsmith, 2002)

V rámci přechodu mezi jednotlivými krátkými – časově malými fázemi projektu je vhodné zabývat se zejména 2 faktory: jaké je další největší hodnota – benefit, který v rámci další etapy dodávám nebo v rámci řešení předávám; a jaká jsou rizika, která afektují pouze další nejbližší fázi projektu (v některých agilních metodikách se též nazývají sprint). Tento princip je tedy též principem prioritizace požadavků a mitigací rizika.

Posledním velkým jmenovatelem v agilních metodikách je spolupráce všech složek v projektu. Bez komunikace není spolupráce. V rámci implementace většiny agilních metodik dochází též k redukování projektových rolí. Je to veliký rozdíl oproti procesním vodopádovým metodikám, které akcentují multi-úrovňové řízení projektu, striktní komunikační a reportingové procesy. Agilní principy zahrnují minimalistické pojetí v projektových rolích a předepsané dokumentaci. Opět se v tomto případě ukazuje, že jejich nasazení do určitých typů projektů (projekty v segmentu státní správy nebo velké dislokované týmy) nebude to pravé.

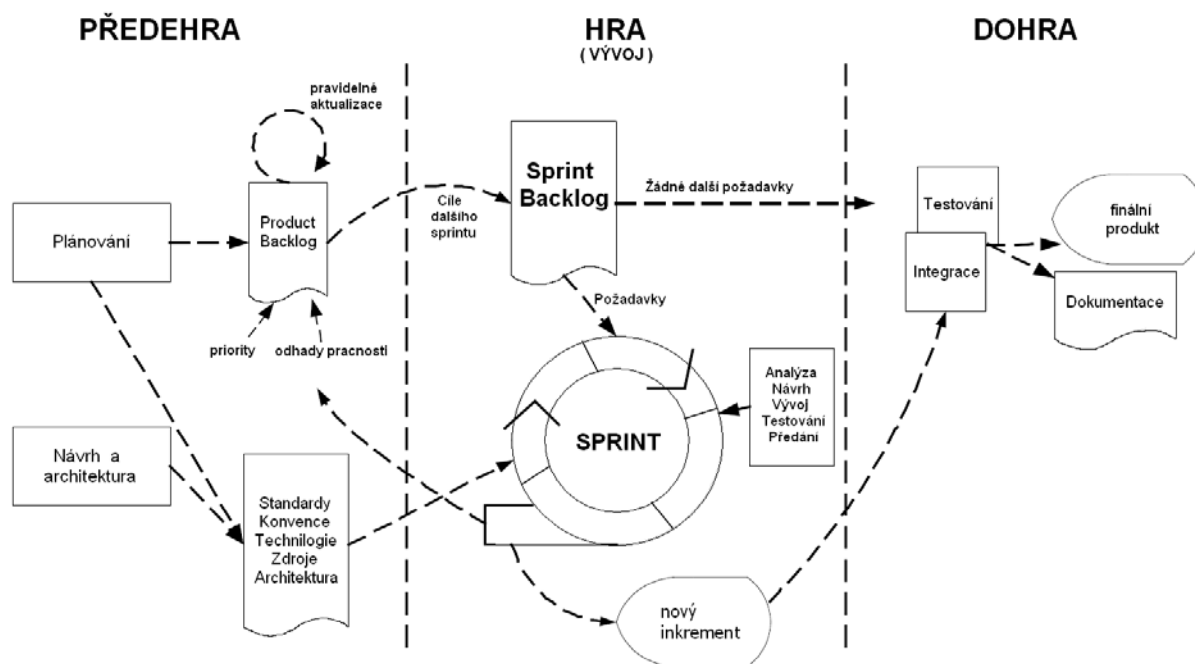
Druhy Agilních metodik

Obecně nejsou agilní metodiky velmi podrobné. Velmi často nepředepisují sáhodlouhý seznam dokumentace nebo procesních modelů. Agilní metodiky dbají na rychlý vývoj dodávek, tak, aby byl zákazník spokojen.

Extrémní programování (XP) je dnes nejznámějším zastupitelem agilních metodik. V době svého vzniku, v roce 1999, se snažilo reagovat na problémy tehdejších projektů a řešit je cestou extrémů. Vychází z principu, který říká, že pokud něco funguje, tak to budeme používat v maximální možné míře (extrémně). Základním pravidlem, které ctí extrémní programování je: „Jediným exaktním, jednoznačným, změřitelným, ověřitelným a nezpochybnitelným zdrojem informací je zdrojový kód“ (Kadlec, 2004). Extrémní programování se tedy zaměřuje především na psaní zdrojového kódu. XP je vhodné na menší a střední projekty, u velkých však již může docházet k chaosu.

SCRUM Metodika byla představena Kenem Schwabem a Mikem Beedlem roku 2002. Přirovnává vývoj SW ke hře ragby. Zaměřuje se na to, že vývoj softwaru je stejně tak jako ragby plně nečekaných událostí a změn a snaží se, aby tým na tyto změny pružně a rychle reagoval. Na rozdíl například od Extrémního programování se zaměřuje spíše na organizační a manažerské aspekty vývoje softwaru, než na psaní samotného kódu.

Nejdůležitějším dokumentem celé metodiky SCRUM je product backlog – do něj se zapisují všechny požadavky zákazníka na výsledný výstup (zpravidla software). Zákazník má u každého požadavku také určit jeho prioritu – může se totiž stát, že některá funkčnost nakonec nebude implementována. Product Backlog by se měl také seřadit podle určitých logických či funkčních celků, aby byl co nejvíce přehledný. Položky Product backlogu jsou průběžně přidávány, odstraňovány a aktualizovány, aby vždy odpovídaly tomu, co je požadováno.



Obrázek 11: SCRUM Metodika

Zdroj: Agilní metodiky (Hajdin, 2005)

Feature Driven Development (FDD) je další z metodik, které dávají pár návrhů a postupů jak řešit především vývoj aplikací. Nezabíhají do velkých podrobností, protože ctí myšlenku, že podrobně specifikované metodiky, které se snaží o zachycení všech kroků vývoje, většinou v praxi selhávají, protože nedokáží průběžně přizpůsobovat vývoj, a jejich pravidla mohou zavést vývoj, který probíhá podle daných pravidel do slepé uličky. Naopak kladou větší nároky na vedoucího projektu, který se nemůže o tyto pravidla opřít a musí řídit vývoj vlastními rozhodnutími a přizpůsobit jej aktuální situaci (Abrahamsson, 2002). Metodiku představili pánové Jeff De Luca a Peter Coad v devadesátých letech minulého století.

Lean development sice není úplně tak metodikou, přináší však sadu pravidel a principů, které se velmi hodí do agilního přístupu vývoje a určitě stojí za zmínku. Vystihující pro tuto sadu pravidel je, že vznikala v Japonsku a původ má v automobilovém průmyslu. Z toho vyplývá mnoho rysů. Japonci moc skromní nejsou, co se cílů týče, a ani Lean Development není. Klade si za úkol pro mnohé neuskutečnitelné cíle:

- vyvíjet software za třetinu obvyklého času,
- vystačit s třetinou obvyklého rozpočtu,
- snížit četnost chyb na třetinu obvyklého množství.

Vznikala v 80. letech, kdy se v Japonsku rozmáhal automobilový průmysl. Japonské automobilky (hlavně Toyota) se snažily co možná nejvíce zjednodušit a zkrátit výrobu aut. Vznikal tzv. „Lean Manufacturing“ a jeho hlavním motem bylo „Absolutní odstranění všeho zbytečného“ (Kadlec, 2004). Postupně se těmito pravidly inspirovali i tvůrci softwaru a aplikovali je na jeho vývoj. Při tomto odstraňování všech zbytečností se vyžaduje, aby každý sebemenší výstup, zpráva, dokument, atd. byl v daném rozsahu podložen hodnotou, kterou přinese výsledné aplikaci, jinak je považován za zbytečný a měl by být odstraněn. Lean Development zavádí, podobně jako další agilní metodiky, také iterativní přístup a zavedení zpětné vazby. Co je možná největší specifikou, zvláště důležitou pro naše zeměpisné šířky, je fakt, že Lean Development klade velký důraz na firemní kulturu a snaží se zavádět co možná největší loajalitu zaměstnanců tolik typickou pro Japonské firmy, kde mají zaměstnanci často jistotu doživotního zaměstnání.

Každá z výše uvedených agilních metodik staví svůj přístup na něčem trochu jiném. Stručný přehled hlavních myšlenek a odlišností sumarizuje tabulka 2.

Tabulka 2: Agilní metodiky

Zdroj: Autor

Metodika	Hlavní myšlenka metodiky	Zaměření metodiky	Nedostatky metodiky
XP	Malé týmy, denní Meziprodukty, párové programování	Důraz na programování	Malý důraz na manažerskou oblast, extrémní praktiky, vhodné jen pro malé týmy
SCRUM	Rychlé, několika denní sprinty, samostatné týmy	Scrum Meetings, zaměření na dodání něčeho nového	Nejsou rozváděny integrační a akceptační oblasti
FDD	Vlastnosti, objektově orientovaná, velmi časté dodávky.	Navržení a implementace systému po vlastnostech.	Zaměřuje se pouze na návrh a implementaci.
LEAN	Odstranění všeho zbytečného	Jde pouze o sadu pravidel. Zaměření se na hodnotu pro zákazníka a kulturu organizace, původem z průmyslu	Důraz na komplexnost organizace
RUP	Kompletní metodika pokrývající celý proces vývoje	Hodně rozsáhlá metodika. Existuje pro ni mnoho doplňkových materiálů a podpůrných nástrojů	Není úplně agilní. Změny musí projít složitější procedurou.

Závěr

Agilní metodiky přináší řešení a hodnotu do řízení ICT projektů. V porovnání s tradičními procesními technikami a metodologiemi, přináší agilní pojetí principy bližší spolupráce, redukce projektových rizik a zvýšené vizibility do průběhu projektu. Zvláště pak v případě takových druhů ICT projektů, jejichž společným jmenovatelem je měnící se prostředí či vyvíjejí požadavky.

Agilní řízení je nepokrytě předmětem dalšího zkoumání a testování na specifických projektech, mající různé atributy (počty požadavků, počet lidí v řešitelském týmu, segment podnikání, atd.). Na základě specifikace agilních principů, definovaných různými agilními metodikami (SCRUM, LEAN, XP), je výstupem této studie fakt, že agilní management nemůže být bez důkladného zvážení jednoduše nasezen do všech druhů ICT projektů. Atributy projektu v korelaci s nasazení agilních metodik by mělo být jasným bodem dalšího zkoumání.

Literatura

Beck K., Cockburn A., Jeffries R., Highsmith J. (2002). *Agile manifesto*, <http://www.agilemanifesto.org>

Boyle, Todd A. (2011). *Learning to be lean: the influence of external information sources in lean improvements*. Journal of Manufacturing Technology Management, ISSN 1741038X

David I. Cleland, Roland Gareis (2006). *Global Project Management Handbook*. McGraw-Hill Professional, 2006. ISBN 0-07-146045-4.

Ernst&Young (2011). *Project Management Practices in the Czech Republic for year 2011*. Research, Ernst&Young, 2011 [cit. 2012-06-08]. Online: <http://www.ey.com/CZ/cs/Issues#publikace>

Fernandez, Daniel J. (2009). *Agile Project Management – Agilism versus Traditional Approaches*. The Journal of Computer Information Systems, p.10-17, v. 49, ISSN 08874417

Goldratt, Eliyahu M. (1997). *Critical Chain*. ISBN 0-88427-153-6

- Hajdin, Tomáš (2005). *Agilní metodiky*. Masarykova univerzita v Brně, Brno.
- Highsmith J., *Agile Software Development Ecosystems*, Addison–Wesley, Boston, MA, 2002.
- Kadlec, Václav (2004). *Agilní programování, Metodiky efektivního vývoje softwaru*. Computer Press, Brno, ISBN 80-251-0342-0
- Kelley, James. *Critical Path Planning and Scheduling: Mathematical Basis*. Operations Research, Vol. 9, No. 3, May–June, 1961
- Kelly Watters, *All about agile*, [http://www.allaboutagile.com/agile-principle-2-agile-development-teams-must-be-empowered/..](http://www.allaboutagile.com/agile-principle-2-agile-development-teams-must-be-empowered/)
- Milosevic, Dragan Z. (2003). *Project Management ToolBox: Tools and Techniques for the Practicing Project Manager*. Wiley. ISBN 978-0-471-20822-8.
- OGC (2009). *Managing successful projects with PRINCE2*. Office of Government Commerce. ISBN 978-0-11-331059-3
- Pekka Abrahamsson, Outi Salo, Jussi Ronkainen & Juhani Warsta (2002). *Agile software development methods – Review and analysis*. VTT Electronics, University of Oulu, ISBN 951-38-6010-8 (URL: <http://www.inf.vtt.fi/pdf/>)
- PMBOK, A Guide to the Project Management Body of Knowledge (2004). Third Edition. Project Management Institute, Pennsylvania USA, ANSI/PMI 99-001-2004, ISBN 1-930699-45-X
- PMI (2013). *2013 Pulse of the Profession – Identifies the High Cost of Low Performance*, dostupné z <http://www.pmi.org/Knowledge-Center/Pulse/Pulse-Key-Findings.aspx>
- Standish Group (2004) Revised from Third Quarter 2004. 2004 Quarterly Reports [online]. The Standish Group. Available at http://www.standishgroup.com/sample_research/PDFpages/q3-spotlight.pdf [accessed 18/11/2005].
- Young-Hoon Kwak (2005), *A brief History of Project Management*. In: The story of managing projects. Elias G. Carayannis et al. (9 eds), Greenwood Publishing Group, 2005. ISBN 1-56720-506-2
- JEL Classification:** M15, O20.

Summary

The Benefits of Agile Project Management Practices in ICT Projects

Agile is bringing solution and values to the various projects, characterized by a large degree of ignorance in the total requirements in a changing environment. In comparison with traditional methods, agile gives clear instructions, clearly functioning in the real project practice, leading to increased visibility, adaptability, and business value and decreasing of the project risks.

Agile management is certainly subject to further examination and testing for specific projects at specific delivery. Without a doubt, agile management can not be simply adjusted to all kinds of IT projects without deeper consideration and research based on exact data coming from particular projects.

Key words: Project Management, Agile, Methodology, Agile Principles.

Využití metodiky COBIT 4.1 při penetračním testování bezpečnosti IS

Tomáš Klíma

xklit10@vse.cz

Doktorand oboru aplikovaná informatika

Školitel: Doc. Ing. Prokop Toman, CSc., (toman@vse.cz)

Abstrakt: Penetrační testy, jakožto efektivní nástroj ověření reálné bezpečnosti informačních systémů, v současné době provází několik problémů, které mohou mít za následek jejich nekvalitní, nekomplexní, a v krajním případě i destruktivní provedení. Mezi ty nejpalčivější patří nedostatečné začlenění řízení penetračního testu do kontextu řízení podnikové informatiky a nedostatečná kvalita dostupných metodik penetračního testování, zejména jejich nekomplexnost a neaktuálnost. Testy jsou následkem toho špatně plánovány (zvláště zaměření a rozsah) a jejich průběh není formálně sledován. Výstupy, včetně doporučení z testu plynoucích, jsou využívány neefektivně – jejich implementace není řádně řízena a dochází tak ke snížení potenciálu zlepšení, které penetrační test mohl organizaci přinést.

Článek představuje možné řešení těchto problémů, které spočívá ve využití autorem nově vytvářené metodiky PETA k ověření vybraných kontrolních cílů metodiky COBIT 4.1, což umožní dosáhnout komplexity testování bezpečnosti IS a zároveň vysoké efektivity při implementaci doporučení z penetračních testů plynoucích. Toto propojení jednotlivých částí metodiky s kontrolními cíli umožňuje zainteresovaným stranám také lépe pochopit význam, rozsah a přínos penetračního testování k ověřování a následně i zvyšování úrovně bezpečnosti informačního systému organizace.

Klíčová slova: Bezpečnost informací, penetrační test, Cobit, PETA

Úvod

V současné době jsou podnikové informační systémy více než kdy dříve ohroženy narušiteli informační bezpečnosti, ať už jde o hackery, nepoctivé zaměstnance či obchodní partnery¹. V souvislosti se stále častějšími bezpečnostními incidenty byl dokonce rok 2011 označován odbornou komunitou jako „year of data breach“, rok 2012 jako o „year of hacktivism“ a právě uplynulý rok 2013 jako o „year of surveillance“.

Důsledkem toho se zdá, že současné metody ochrany informačních systémů ztrácejí na své efektivitě, když i takové organizace jako je RSA, Global Payments, nebo LinkedIn zažily úniky dat, které měly pro některé z nich téměř fatální dopad (Check Point, 2013). Nemůže být pochyb, že tyto subjekty měly všechna tradiční protiopatření (firewall, IDS/IPS, monitoring síťových toků, antivir a mnoho dalších) v době incidentu nasazena a řádně konfigurována, přesto však byla útočníkem nalezena zranitelnost, která umožnila kompromitaci informačního systému.

Jedním z možných řešení může být větší zaměření na ofenzivní techniky používané při zajišťování informační bezpečnosti.^{2,3} Jde zejména o penetrační testování (také nazývané jako etický hacking), které využívá ty samé nástroje a postupy jako útočníci za účelem odhalení maxima zranitelností, které následně organizace ať už vlastními silami, či pomocí třetí strany odstraní. Účelem penetračního testu je ověřit úroveň realizace bezpečnostních politik technickými a dalšími prostředky, a z pohledu útočníka kompromitovat cílový informační systém a identifikovat možné dopady reálného útoku. Jediným pod-

¹ Přičemž dle (Verizon, 2013) je původcem 92% datových úniků externí pachatel (tedy hacker).

² S čímž dle výzkumu (Ernst, 2013) souhlasí 69% dotázaných společností, které penetrační testy uvádí jako jednu z top 4 priorit v oblasti informační bezpečnosti.

³ V souladu s (Mahmood, 2010), který uvádí, že výzkum by v budoucnosti měl být významněji zaměřen na techniky využívané útočníky, oproti stávající praxi, jež se zaměřuje primárně na výzkum obranných mechanismů.

statným rozdílem mezi reálným útočníkem (black hat hacker) a testerem (white hat neboli ethical hacker) je jejich motivace⁴. Útočník se snaží průnikem získat jistý prospěch (nebo alespoň uškodit organizaci), kdežto tester se snaží odhalením co nejvíce bezpečnostních slabin prospěch organizaci přinést. Málokdy je ovšem tester jedinou osobou zainteresovanou v celém procesu, jelikož zvláště v případě rozsáhlejších testů bývá na straně dodavatele zapojen celý tým (také nazývaný red team, neboli útočný) a na straně zadavatele je také nejméně jeden pracovník, který celý test poptává a následně řídí jeho průběh i odstranění zjištěných zranitelností. Pochopitelně pokud si organizace test zajišťuje vlastními silami, což bývá případ zejména těch větších, struktura celého týmu a jejich vztahy se mění.

V dosažení maximálního přínosu testu pro organizaci ovšem oběma stranám smluvního vztahu v současné době brání několik problémů. Těmi nejpalčivějšími jsou nedostatečné začlenění řízení penetračního testu do kontextu řízení podnikové informatiky a nedostatečná kvalita dostupných metodik penetračního testování, hlavně jejich nekomplexnost a neaktuálnost. To má za následek stav, kdy jsou testy špatně plánovány, zejména jejich zaměření a rozsah. Také jejich průběh není formálně sledován a výstupy, včetně doporučení z testu plynoucích, jsou využívány neefektivně a implementace nápravných opatření není řádně řízena.

Tyto dva hlavní problémy je možné buďto řešit odděleně, tedy sestavit novou metodiku, která shrne nejnovější poznatky a postupy do jediného rámce a následně pak hledat souvislosti mezi řízením penetračního testu (včetně jeho projektování) a řízením podnikového IT, nebo je možné k problémům přistoupit komplexně a využít synergie při provázání nově vyvíjené metodiky s rámcem, který nabízí některá z rozšířených metodik řízení podnikového IT. Autor tedy dále pracuje s druhým přístupem a představí nejen nově vyvíjenou metodiku PETA, ale následně nastíní i její praktické propojení s kontrolními cíli rámce řízení podnikového IT COBIT 4.1.

Metodika penetračního testování PETA

Nedostatky v současně používaných metodikách penetračního testování vedou firmy (či testovací týmy) nabízející takovéto služby k vytváření vlastních postupů, které ovšem málokdy mají pevnou strukturu a většinou jsou více či méně závislé na tacitních znalostech konkrétních testerů, které s jejich odchodem následně z organizace mizí. Také publikované návody na testování, či „hackování“ konkrétních platform či zařízení (Selecký, 2012), (Scambray, 2009), mohou být užitečné jako podkladový materiál, ale zpravidla jim chybí pevnější rámec a tudíž mají smysl jen pro odborníka, který je dokáže použít v kontextu celého testovacího procesu. Testeré sice často uvádějí využívání klasických metodik, jakými jsou ISSAF (OISSG, 2006) či OSSTMM (Herzog, 2006), v praxi se jimi ale kvůli zastaralosti neřídí⁵. To všechno ilustruje současný stav, kdy se dá říci, že neexistuje metodika penetračního testování, která by byla jak komplexní, tak obsahovala aktuální postupy a byla reálně využitelná při penetračním testu.

S tímto problémem se autor vypořádává vytvořením vlastní metodiky s názvem PETA, jejíž základy byly položeny v roce 2011 v souvislosti s penetračním testováním bezdrátových sítí (Klíma, 2011). Od té doby byla postupně rozšiřována pro použití v rámci testování kompletní infrastruktury.

Metodika je rozdělena do následujících čtyř úrovní:

1. Obecný postup
2. Detailní postup
3. Speciální postupy
4. Nástroje

První úroveň je nazvána obecná a prezentuje hrubé rozdělení činností a projektování penetračních testů. Druhá, detailní, následně definuje podrobný rozpad činností až na úroveň jednotlivých kroků vedoucích od prvotního kontaktu s infrastrukturou až k jejímu, pokud možno úplnému, ovládnutí. Třetí úroveň obsahuje postupy testování konkrétních prvků IS a poslední, čtvrtá, pak popis nástrojů využitých při

⁴ Druhým rozdílem je odlišný přístup auditora a hackera ve smyslu systematickosti přístupu. Hacker obvykle nepostupuje podle dané metodiky, spíše vyzkouší několik útoků a v případě neúspěchu volí jiný způsob získání dat. Pokud však hacker zná a používá dále představenou metodiku, i jeho postup může být systematický.

⁵ Jako další z mnoha často uváděných relevantních publikací je možné jmenovat (SANS, 2010) a (NIST, 2008).

testu. Toto dělení je použito hlavně proto, že první dvě úrovně mohou být použity i v dlouhodobém horizontu bez ztráty aktuálnosti a pouze třetí a čtvrtou je třeba v budoucnosti aktualizovat.

Pro snadnější provázání metodiky PETA s rámcem COBIT 4.1 jsou v tabulce 1 uvedeny kroky spadající do detailního postupu, logicky členěny do fází plánování, testování, reporting. Tyto kroky jsou následně podrobněji představeny. V tabulce 2 jsou pak uvedeny speciální postupy popisující testování jednotlivých prvků IS.

Tabulka 2: Detailní postup

Zdroj: autor

Plánování	Testování	Reporting
1.1) Identifikace potřeby	2.1) Sběr informací	3.1) Úklid testovaného prostředí
1.2) Identifikace stakeholderů	2.2) Mapování sítě	3.2) Analýza dokumentů
1.3) Sestavení zadavatelského týmu	2.3) Identifikace zranitelností	3.3) Vytvoření reportu
1.4) Stanovení rozsahu testu	2.4) Penetrace	3.4) Prezentace reportu
1.5) Stanovení pravidel testu	2.5) Získání přístupu a eskalace privilegií	
1.6) Podání poptávky	2.6) Kompromitace IS	
1.7) Vyhodnocení nabídek	2.7) Udržení přístupu	
1.8) Sepsání smlouvy	2.8) Zahlázení stop	
1.9) Kick off meeting		

V rámci metodiky je nutné nejprve se, ve fázi plánování, zaměřit na identifikaci potřeby, kterou může být nutnost dodržet určitý standard (např. PCI DSS⁶), dlouhodobý plán⁷, nebo úkol otestovat nově zaváděnou součást IS. Dále je třeba identifikovat všechny nositele zájmu (stakeholdery) a sestavit zadavatelský tým. Zadavatelský tým následně vybírá tým řešitelský, ať už se jedná o interní zdroje, nebo jde o řešitele externího. Současně je stanoveno zadání, které určuje jak rozsah (scope), tak pravidla samotného testu⁸. Zejména je dohodnuto, jaké nástroje a postupy testeré smí používat a jaká úroveň kompromitace infrastruktury je žádoucí. Pokud jsou tedy řešitelský tým i zadání sestaveny, nic nebrání zahájit samotný test (kick off, neboli zahajovací setkání).

V rámci testu nejprve tester sbírá maximum informací o cílové infrastruktuře, většinou z veřejně dostupných zdrojů. V této fázi ještě zpravidla se samotnou infrastrukturou nepřichází do styku. V momentě, kdy má dostatek informací, přechází k samotnému skenování a mapování sítě (infrastruktury). Od této chvíle již může být správci IS detekován, proto přesné provedení a intenzita mapování závisí na typu testu (pokud není test správcům nahlášen, tester zpravidla snižuje intenzitu skenování a snaží se zůstat nepozorován). Po zmapování sítě má již představu o její topologii a běžících službách, proto začíná hledat zranitelné místo, které mu následně umožní průnik do IS⁹. Pokud se mu průnik podaří, jeho cílem je získat na daném prvku IS maximální možné oprávnění, které mu následně umožní rozšířit pole působení. Právě kompromitace co největší části infrastruktury bývá častým cílem a záleží jen na stanovených pravidlech testu a na kvalitě zabezpečení, kde se tester ve svém postupu zastaví. Po získání prvotního přístupu se snaží si zajistit přístup trvalý, ať už ve formě založení nového uživatelského účtu či instalace služby umožňující vzdálený přístup. Pokud je to součástí zadání, po dokončení předchozích kroků zakryje stopy své přítomnosti v auditních a monitorovacích systémech.

Tím je samotný test považován za ukončený a tester už jen, po sestavení evidence své přítomnosti v systémech, má za povinnost uvést IS do stejného stavu, v jakém do něj vstupoval, zejména ho vyčistit od všech nástrojů a služeb, které využíval, a potencionálně by snižovaly jeho další bezpečnost.

⁶ Payment Card Industry Data Security Standard, www.pcisecuritystandards.org

⁷ Např. často bývá managementem požadován interní penetrační test každé dva roky.

⁸ Zadání může stanovovat buď pouze zadavatelský tým, nebo, jak je v praxi běžnější, stanoví pouze hrubé požadavky a konkrétní detaily doplní na základě konzultace s vybraným řešitelem.

⁹ Zde je nutné zmínit, že přesný postup závisí na typu testu (interní/externí), rozsahu (komplexní test/test vybraných prvků IS) a dalších faktorech. V tomto článku je tedy uveden pouze nejběžnější „modelový“ postup.

Poté testera čeká již jen analýza sesbírané dokumentace a evidence o zjištěných slabínách, na základě které pak sestaví závěrečnou zprávu, kterou po akceptaci zadavatelem prezentuje vybranému okruhu pracovníků. Po skartování (nebo alespoň anonymizaci) dokumentace je jeho práce u konce.

Tabulka 2: Speciální postupy

Zdroj: autor

Speciální postupy	Zkratka
Webové aplikace	WA
Webové servery	WS
Sociální inženýrství	SI
Malware	MW
DOS, DDOS	DOS
Sniffing	SNI
Bezdrátové sítě	WI
Fyzická bezpečnost	FB
Koncové stanice	EP
Testování detekce a incident response	IR
Mobilní zařízení	MD

Dodržení zmíněného detailního postupu umožní systematický a metodicky správný postup testu. Pokud není cílem komplexní otestování IT infrastruktury, ale rozsah je omezen, je možné ze speciálních postupů vybrat ty relevantní. Z logiky věci vyplývá, že při takovémto omezení bude i detailní postup aplikován v omezené podobě, zejména ve fázích 2.1 – 2.8.

Ověření vybraných cílů COBITu 4.1 pomocí metodiky PETA

Z portfolia kontrolních cílů obsažených v COBITu 4.1 je nutné vybrat relevantní položky ověřitelné penetračním testem, čímž vznikne zúžený rámec. Definováním takového rámce jsou zároveň specifikovány oblasti, na které se test může zaměřit, což pomáhá zejména ve fázi plánování. Výstupy a reporty z testu jsou pak použitelné jako podklad pro zvyšování úrovně zabezpečení daných oblastí a pro identifikaci rozdílů oproti žádoucímu stavu (gap analysis). Zároveň dochází k odstranění velkého problému, se kterým se setkávají zadavatelé testu – neefektivní, či nesprávná implementace závěrů a doporučení, která vede k znehodnocení možného přínosu pro bezpečnost a kvalitu informačního systému. Propojení s kontrolními kritérii umožňuje jednotlivé závěry vztáhnout k určitým oblastem řízení IS a delegovat a řídit proces jejich implementace, jak v krátkém, tak dlouhém časovém horizontu s ohledem na jejich prioritu. Také umožňuje stanovit oblasti řízení IT, na které by se měla zaměřit pozornost managementu (nejen) z hlediska bezpečnosti.

Každý kontrolní cíl lze ověřit pomocí jednoho či více detailních postupů a jednoho či více speciálních postupů (viz tabulka 3). Je to dáno tím, že například při ohodnocení rizik (cíle P09) je zapotřebí provést test v plné jeho šířce za účelem identifikace zranitelností (a jejich závažnosti) v celé infrastruktuře. Naopak při ověřování reakce správců na incident (cíle DS8), kterým může například být průnik do infrastruktury, jde zejména o ověření kvality incident response procesu, tedy o body 2.2 - 2.8 detailního postupu. Co se týká speciálních postupů, tak každý z nich může být spouštěcím mechanismem pro detekci a případnou eskalaci incidentu, ale relevantní je pouze Testování detekce a incident response (IR).

Dalším příkladem může být cíle AI2 Acquire and maintain application sw. K jeho ověření dochází v krocích 2.3 – 2.7, tedy od detekce slabín použitelných k průniku až po udržení přístupu na ovládnutý systém. Pokud je proces patchování a konfigurace aplikačního softwaru v pořádku, je pro testera těžké stanovit vektor útoku a i eskalace privilegií, či udržení přístupu nemusí být jednoduché. Naproti tomu špatně spravovaná infrastruktura s neaktuálním softwarem je přímou pozvánkou k jejímu rychlému ovládnutí.

Jinými slovy nám tedy toto mapování říká, jaké procesy a oblasti řízení podnikového IT můžeme pomocí penetračního testu ověřit, a jaké postupy, ať už detailní nebo speciální (dle terminologie metodiky PETA) k tomuto ověření použít.

Tabulka 3: Kontrolní cíle

Zdroj: autor

Kontrolní cíl	Detailní postup	Speciální postupy ¹⁰
Plan and organise		
P09 Assess and manage IT risks	1.1 - 3.4	Všechny
Acquire and implement		
AI1 Identify automated solutions	2.2 – 2.5	WA, WS, MW, DOS, SNI, WI, FB, EP, IR, MD
AI2 Acquire and maintain application sw	2.3 – 2.7	WA, MW, DOS, SNI, EP, IR, MD
AI3 Acquire and maintain technology infrastructure	2.2 – 2.7	WA, WS, MW, DOS, SNI, WI, EP, IR, MD
AI6 Manage changes	1.1 - 3.4	Všechny
AI7 Install and accredit solutions and changes	1.1 - 3.4	Dle scope testu
Deliver and support		
DS4 Ensure continuous service	1.1 - 3.4	DOS, IR
DS5 Ensure systems security	1.1 - 3.4	Všechny
DS8 Manage service desk and incidents	2.2 - 2.8	IR
DS9 Manage the configuration	1.1 - 3.4	WA, WS, DOS, SNI, WI, EP, MD
Monitor and evaluate		
ME2 Monitor and evaluate internal control	1.1 - 3.4	Všechny
ME3 Ensure compliance with external requirements	1.1 - 3.4	Všechny, případně dle oblasti compliance

Využití v praxi

V praxi již bylo využití kontrolních cílů COBITu 4.1 částečně ověřeno při řízení implementace doporučení z penetračního testu plynoucích. Na základě přiřazení nálezů (a relevantních doporučení) k jednotlivým kontrolním cílům a následně k zodpovědným osobám bylo možné posoudit stav bezpečnosti informačního systému a identifikovat, který oblastem je třeba věnovat pozornost. Zároveň byl sledován průběh odstranění jednotlivých zranitelností (nálezů).

Ukázka dokumentace jednoho nálezu (kterých mohou být desítky) je uvedena v tabulce 4. Webový server s identifikací AAA má nález typu zranitelnost (další jsou pak informace, průnik a pozitivní zjištění, ale mohou být i další), kterou je dostupnost SSL ve verzi 2.0, zjištění se týká portu 443 (tedy https). Doporučení zní zakázat šifrování pomocí SSL 2.0 a povolit jen novější a bezpečnější SSL 3.0. Jelikož za konfiguraci tohoto serveru je zodpovědný pan Petr Novák, byl mu stanoven úkol zajistit trvalé řešení tohoto problému – dočasné řešení se používá pouze tehdy, existují-li provozní důvody, které brání aplikaci trvalého řešení v rozumném časovém horizontu (jedná se tedy o „workaround“). Před vydáním tohoto úkolu ještě byla provedena analýza, která ověřila právě dopad na provozní prostředí (např. možnost nedostupnosti serveru na klientech se starším programovým vybavením). Sloupec stav

¹⁰ Zkratky viz tabulka 2, speciální postupy.

plní funkci semaforu, který indikuje, jak daleko je řešení problému a jaká je jeho závažnost. Nález je zařazen pod kontrolní cíl DS9 Manage the configuration.

Tímto byla tedy ilustrována praktická použitelnost přístupu při řízení implementace doporučení. Využitelnost při plánování a samotném řízení (a provádění) testu autor plánuje ověřit v rámci své disertační práce, jejíž část bude věnována právě praktickému experimentu, při kterém budou testovány pomocí metodiky PETA jak IS nasimulované ve virtuálním prostředí, tak IS reálně běžící.

Tabulka 4: Implementace doporučení

Zdroj: autor

Pentest - implementace doporučení										
ID	Zařízení	Port	Typ nálezu	Doporučení	Dočasné řešení	Vlastník	Trvalé řešení	Vlastník	Pozn.	Stav
Deliver and support - DS9 Manage the configuration										
1	Web server AAA	TCP 443	Zranitelnost	Zakázat šifrovat komunikaci použitím protokolu SSL v2.	Není	Není	Bude zakázáno k 1.1.2014, dále jen SSL 3.0	Petr Novák	Provedena analýza dopadu na provoz – schváleno.	
2	...	...	...	...	...	...	...	...	...	

Závěr

Jako hlavní problémy penetračního testování bylo identifikováno nedostatečné začlenění řízení penetračních testů do kontextu řízení podnikové informatiky a nedostatečná kvalita dostupných metodik penetračního testování. V článku byl představen přístup, v rámci něhož autor aplikuje jím vytvářenou metodiku PETA k ověření vybraných kontrolních cílů COBITu 4.1, které tvoří tzv. zúžený rámec. Tím směřuje k řešení obou problémů najednou, což by mělo přinést zlepšení v podobě efektivnějšího a komplexnějšího řízení penetračních testů, včetně jejich plánování a řízení implementace doporučení z testu plynoucích. Právě na praktickém řízení implementace doporučení byla ověřena funkčnost tohoto přístupu a bylo pozorováno, že oproti předchozím testům, kdy byly úkoly přiřazovány a kontrolovány ad hoc, došlo k jasnějšímu zlepšení komunikace mezi zainteresovanými osobami, k lepšímu sledování a vyhodnocování plnění úkolů z doporučení plynoucích a zejména díky propojení s jednotlivými kontrolními cíli (a zároveň i firemními procesy) k mnohem jasnějšímu přehledu o úrovni bezpečnosti celého informačního systému. Tím byly zároveň identifikovány oblasti, kam by primárně měli pracovníci bezpečnosti zaměřit pozornost. Naznačené praktické využití potvrdilo, že přístup přispěl k pevnějšímu začlenění řízení penetračních testů (respektive části implementace doporučení) do kontextu řízení podnikové informatiky.

Literatura

ERNST & YOUNG. Global information security survey 2013. Londýn, 2013.

HERZOG, Peter. Open-Source Security Testing Methodology Manual [online]. [s. l.]: ISECOM, 2006. Dostupné z WWW: <<http://www.isecom.org/osstmm/>>.

CHECK POINT. Check Point 2013 security report. Israel, 2013.

IT GOVERNANCE INSTITUTE. Cobit 4. 1. 2007.

KLÍMA, Tomáš. HODNOCENÍ BEZPEČNOSTI BEZDRÁTOVÝCH SÍTÍ. Praha, 2011. Diplomová práce. KIT VŠE. Vedoucí práce L. Pavlíček.

MAHMOOD, M. Adam. MOVING TOWARD BLACK HAT RESEARCH IN INFORMATION SYSTEMS SECURITY. MIS Quarterly. 2010, roč. 3, č. 34.

NIST. Technical Guide to Information Security Testing and Assessment: NIST Special Publication 800-115. 2008.

OPEN INFORMATION SYSTEMS SECURITY GROUP. Information systems security assessment framework [online]. [s. l.]: Open information systems security group, 2006. Dostupné z WWW: <<http://www.oisssg.org/downloads/issaf-0.2/index.php>>.

SANS INSTITUTE. Penetration testing in the financial services industry. 2010.

SCAMBRAY, Joel; MCCLURE, Stuart: Hacking exposed, New York 2009, McGraw-Hill, ISBN: 978-0-07-161375-0.

SELECKÝ, Matúš. Penetrační testy a exploitace. 1. vyd. Brno: Computer Press, 2012, 303 s. ISBN 978-80-251-3752-9.

VERIZON. 2013 Data breach investigation report. New York, 2013.

JEL Classification: L80, L86, M10, M15

Summary

The use of methodology COBIT 4.1 for IS penetration testing

Penetration tests as an effective tool for verification of the real security of information systems currently suffer from several problems that can result in poor quality, non-complex, and in extreme cases even destructive test result. One of the main problems is the lack of integration management penetration test in the context of business ICT and insufficient quality of available penetration testing methodologies. This results in poorly planned tests, in particular their scoping. Their progress isn't formally monitored and test results, including recommendations coming from the test are used inefficiently - their implementation is not properly managed.

The article presents a possible solution to these problems - to use the author's new methodology PETA to verify the selected control objectives of COBIT 4.1, which allows us to achieve high complexity of security testing of IS and also achieve high efficiency in the implementation of the conclusions and recommendations of penetration tests. This methodology - control objectives mapping allows interested parties to better understand the meaning, scope and benefits of penetration testing for verification and consequently improvement of security of information system.

Keywords: Information security, penetration testing, Cobit, PETA.

Alokace lidských zdrojů v projektovém řízení

Roman Nedzelský

xnedr04@vse.cz

Doktorand oboru aplikovaná informatika

Školitel: prof. Ing. Jaroslav Jandoš, CSc., (jandos@vse.cz)

Abstrakt: V současnosti se ve většině podnicích projekty řídí pomocí různých podpůrných nástrojů. Velká část firem využívá různé, již osvědčené mechanismy na tvorbu časových odhadů a na alokaci lidských zdrojů k jednotlivým úkolům. Žádný ze softwarových nástrojů ovšem není natolik sofistikovaný, aby řešil jeden z nejčastějších problémů dnešní doby - alokaci lidských zdrojů - na pokročilé úrovni.

Práce se věnuje jak standardním způsobům alokace lidských zdrojů, tak alokaci s ohledem na okolní ad-hoc vlivy a tuto problematiku přenáší do prostředí softwarových nástrojů na podporu projektového řízení. Zabývá se hlavními nedostatky procesu alokace lidských zdrojů a možnými dopady na výsledek jednoho, případně více souběžných projektů.

K řešení stávajících nedostatků autor doporučuje využití dalších datových zdrojů, které by dodaly potřebné informace do procesu plánování (případně do softwarového nástroje na podporu projektového řízení) a navazující úpravu algoritmu pro alokaci a vyvažování lidských zdrojů.

Klíčová slova: Enterprise Project Management, Resource Management, PERT, Monte Carlo, IS Integration.

Úvod

V rámci řízení IT projektů bývá velká jejich část neúspěšná. To je zapříčiněno mnoha důvody (viz Obrázek 1). Jednou z takových příčin je i alokace lidských zdrojů, na niž se zaměřuje tento příspěvek. Na Obrázku 1 je možné vidět tuto oblast v částech grafu „skill issues“, případně zčásti i v „Execution Issues“. V multi-projektovém prostředí dle Blocha a kol.(2012) končí polovina projektů s náklady nad 15 milionů dolarů z 45% s vyššími náklady, z 7% delší dobou trvání a z 56% s nižší dodanou funkcionalitou oproti naplánovaným hodnotám.

Rough distribution by cause of the 45% of IT projects that experience cost overruns (for those with budgets >\$15 million in 2010 dollars), %



IT projects with budgets >\$15 million

Cost overrun, 45%

Schedule overrun, 7%

Benefits shortfall, -56%

Source: McKinsey-Oxford study on reference-class forecasting for IT projects

Obrázek 12 - Úspěšnost IT projektů

Zdroj: McKinsey (2013)

V rámci projektového řízení je možné provádět plánování jednotlivých úkolů WBS¹ a zdrojů k nim přiřazeným několika způsoby. Lze využít „osvědčené“ metody, jakými jsou například PERT², Monte Carlo nebo CPM³. Tyto metody již ovšem patří k těm zastaralejším. Trendem poslední doby jsou především agilní metodiky projektového řízení a plánování. Jednu z nejobtížnějších částí samotného plánování v multi-projektovém prostředí pak představuje vhodný výběr z velkého množství potencionálních lidských zdrojů. V momentu samotného plánování je nutné udržovat aktuální informace o tom, jak moc je konkrétní zdroj na souběžných projektech alokovaný, podle jakého kalendáře se řídí a jestli nemá naplánované nějaké mimofiremní aktivity, jakými může být například dovolená, návštěva lékaře, realokace do jiné země či na jiný projekt apod.

Hlavní cíle tohoto příspěvku jsou následující:

- Provést analýzu problémů v oblasti alokování lidských zdrojů v multi-projektovém prostředí a definovat základní příčiny těchto problémů – tedy nalézt problémy spojené s alokací lidských zdrojů.
- Navrhnout koncept řešení jednotlivých problémů identifikovaných v předešlém bodu.
- Navrhnout doporučení, jak by tyto problémy mohly být řešeny za použití modifikovaných algoritmů v rámci vybraného softwarového nástroje.

Obecný postup plánování projektu

Obecný postup v plánování projektu představuje rozhodně nejznámější a nejpraktikovanější postup ve většině společností, které používají projektové řízení. Na jeho začátku jsou sepsány jednotlivé úkoly projektu v hierarchické podobě, k nimž se definuje délka jejich trvání. Následně resource manager, případně projektový manažer, přiřadí k úkolům jednotlivé zdroje (lidské, materiální i nákladové). Tak je sestaven projektový tým, který je pak v rámci projektů rozvíjen a řízen. Výběr vhodných zdrojů představuje důležitou část celé plánovací fáze projektu. V této fázi je nutné zvážit mnoho různých aspektů, a proto zde často participuje i resource manager. Mezi vlastnosti jednotlivých lidských zdrojů lze například řadit jejich hodinovou sazbu (kterou je nutné zvážit především kvůli rozpočtu projektu), dovednosti zdroje (jeho skill-set), RBS⁴ zdroje, tým, oddělení a případně zemi, ve které zdroj pracuje. Tento obecný postup se nejsnáze praktikuje v menších společnostech, případně v menších týmech, kde je maximálně několik desítek lidských zdrojů se snadno zjistitelným vytížením. Poté musí projektový manažer rozhodnout, zda bude k jednotlivým úkolům přičítat dobu trvání vynásobenou předem definovaným koeficientem, aby vytvořil tzv. projektové nárazníky na projektu a zajistil tak případnou rezervu pro špatně odhadnuté úkoly a pojistil se před nepříznivými okolními ad-hoc vlivy. Projektový manažer by měl mít také v rámci týmu dostatečnou autoritu a měl by se ujistit, že každý z jednotlivých členů projektového týmu je seznámen s obsahem jemu přidělených úkolů, že těmto úkolům rozumí a že souhlasí s tím, že je schopen je splnit (Bowen, 2012; Hart, 2012; Raunak & Osterweil, 2005; Roberto, Bruno, & Jose-Luis, 1976).

Mimo standardní odhadnutí časového rámce plánovaného úkolu (tedy odhadu na základě provedené analýzy, vlastních zkušeností, případně zkušeností senior architekta řešení apod.) je možné přibližnou dobu jeho trvání stanovit i pomocí jiných technik. Jednou z nich je například metoda Monte Carlo. Ta patří mezi složitější techniky odhadu doby trvání úkolů a hojně se používá i pro stanovení pravděpodobnosti stanoveného projektového rizika. Je založena na využití náhodných veličin a teorii pravděpodobnosti. V případě odhadu doby trvání je celý postup založen na pesimistickém a optimistickém odhadu (mantinelech). Na tyto hodnoty (pro jednotlivé úkoly/souhrnné úkoly) je poté aplikována deterministická analýza, přičemž tento proces je opakován několikrát, nebo i několikrát. Nevýhodou je tedy její časová náročnost (z důvodů statisticky vypovídajícího souhrnu dat), což má za následek její ne příliš časté používání v praxi (Dvořák & Lichtenberg Lisý, 2013).

¹ WBS (Work Breakdown Structure) – technika v projektovém řízení, kdy je dodávka rozložena do po sobě jdoucích dílčích dodávek.

² PERT – analýza založená na výpočtu doby trvání na základě tří odhadů a následného výpočtu váženého průměru.

³ CPM (Critical path method) – metoda kritické cesty – jedná se o sekvenci vzájemně závislých kroků, na nichž závisí úspěch projektu (z pohledu času, i rozsahu).

⁴ Resource Breakdown Structure, tedy zařazení lidského zdroje v hierarchii společnosti.

Další možností stanovení doby trvání úkolu je použití PERT analýzy, jejíž vznik lze datovat do období 2. sv. války. Využívá principu stanovení optimistické, realistické a pesimistické doby trvání úkolu. Z těchto tří údajů se pak prostřednictvím vzorce stanoví doporučená doba trvání úkolu. Tuto analýzu obsahovalo původně mnoho softwarových nástrojů na řízení projektů (později z nich vymizela a objevuje se zde nyní jen v podobě neoficiálních doplňků), protože není tak výpočetně náročná jako Monte Carlo a v současnosti je před Monte Carlo upřednostňovaná.

V rámci plánování projektu je důležité především stanovení priorit jednotlivých aktivit, případně i celých projektů v rámci multi-projektového prostředí. Priority lze nejjednodušeji nastavit prostřednictvím stanovení číselné hodnoty jednotlivým úkolům nebo projektům. Na základě tohoto údaje se pak může provádět vyvažování zdrojů v rámci daného programu nebo projektového portfolia. Rozšířením takového postupu prioritizace může být stanovení výpočetního algoritmu priorit z různých vstupních informací. Vyvažování zdrojů je nutné provádět především v organizaci s vyšším počtem projektů, kde zdroje obecně (tedy nejen zdroje lidské) participují na více projektech a jsou tedy mezi projekty sdílené. Takové vyvažování se nejčastěji provádí ručně, tedy na základě domluvy mezi projektovým manažerem a resource manažerem, který má jednotlivé zdroje na starosti.

Z pohledu HRM⁵ se dá konstatovat, že by HR⁶ oddělení mělo neustále rozvíjet lidské zdroje a zkvalitňovat tak lidský kapitál, který by vedl k větší efektivitě, a tím i navýšení výkonnosti celé firmy. Tím by se zkrátily délky trvání jednotlivých úkolů v rámci projektového řízení a navýšoval by se tak zisk. Rozvoj zdrojů by bylo možné realizovat i tím, že by prostřednictvím shromažďování vhodných informací o zdrojích mohly být jednotlivé lidské zdroje alokovány na úkoly, které by je bavily (v rámci jejich zájmů atd.). Tím by se zvýšila pravděpodobnost efektivního pracovního výkonu, tzn. pravděpodobnost dokončení úkolu/projektu za menší jednotku času, což by vedlo k menším nákladům na zdroj. Z toho tedy vyplývá nejen vyšší efektivita, ale i větší výnos v rámci projektu/projektů.

Problém alokace lidských zdrojů

Na základě analýzy (rešerše) odborné literatury a vlastních zkušeností autora lze specifikovat následující problémy alokace lidských zdrojů ve stávajících přístupech (algoritmech projektového řízení):

- Nedostatek informací pro vhodné navržení lidského zdroje pro konkrétní úkol.

Nedostatek informací lze pokládat za největší nevýhodu v oblasti plánování lidských zdrojů i projektů obecně. V důsledku se jedná se o největší snížení efektivity práce a výdělku na jednotlivých projektech společnosti. Zdroj může být neefektivním z několika důvodů. Nemusí danou problematiku umět, tedy nemá znalosti, nebo problematiku umí, ale je junior a neumí ji na vysoké úrovni, případně mu chybí zkušenosti. Jinou možnou příčinou je, že je zdroj pře-alokován a nestíhá plnit jednotlivé úkoly – potažmo všechny úkoly. Další možný důvod ke snížení efektivity představuje motivace. Pokud je zdroj demotivován, stává se taktéž neefektivním, protože nemá žádný cíl, kterého by chtěl v rámci plnění úkolu dosáhnout.

Lidské zdroje v produktu Microsoft Project (ale i v jiných softwarových produktech – např. Primavera a další.) se vybírají dvěma způsoby. Prvním z nich je výběr konkrétního zdroje ze seznamu sdílených zdrojů organizace. Po nahrání projektového plánu na projektový server je možné provádět optimalizace lidského zdroje (v rámci týmového plánovače, případně vyrovnávání zdroje). Druhý způsob představuje nejprve navržení generických rolí, tzn. přiřazení rolí jednotlivým zdrojům organizace, které se budou jménem shodovat s generickými, a až poté provedení návrhu v nahrazování generických zdrojů konkrétními. Při této možnosti se zobrazuje volná kapacita lidského zdroje, je možné zadat upřesňující údaje do podnikového informačního systému, jako například datum, případně další rozřazení v organizační struktuře firmy apod. (Hart, 2012). Lidské zdroje ovšem nemají definované své vlastnosti/specializace a nejsou pro ně dostupné ani další informace z ostatních podnikových systémů, které by odrážely jejich osobnostní a profesní rozvoj (což by do procesu plánování zanášelo i jistou formu motivace). Takové vyvažování (tedy s omezenými informacemi) může vést ke snížené efektivitě i horšímu průběhu celého projektu (ztráta kvality výstupů, jejich opožděné dodání nebo vyšší náklady na projekt).

⁵ HRM (Human Resource Management) – řízení lidských zdrojů.

⁶ HR (Human Resource) – lidské zdroje.

- **Lidské zdroje na plný, resp. na částečný (např. poloviční) úvazek**

Stávající přístupy u lidských zdrojů typicky nezohledňují jejich kapacity pro projekt, tj. nezohledňují, zda jsou zdroje v dispozici na plný (full-time), nebo částečný úvazek (part-time). Organizace často k vyrovnaní personálního portfolia využívají externí lidské zdroje, které použijí na konkrétní běžící projekty, aby nedošlo k nevyváženosti celého portfolia. Dalším důvodem k najmutí externího zdroje může být ale i nutnost vykonat činnost na projektu, kterou se nikdo ve společnosti doposud nezabýval. Úkol by trval o mnoho déle, než když ho provede externí zdroj s požadovanými znalostmi i zkušenostmi.

Často ale dochází k situacím, kdy jsou externí zdroje nepracující na plný úvazek zahrnovány do dalších projektových plánů nově vznikajících projektů. Mohou nastat i situace, kdy je nutné uvolnění zdrojů z jednoho projektu na jiný v rámci prioritizace. Tyto lidské zdroje jsou zapojeny do více projektů najednou a přesouvají se z jednoho týmu do druhého. Tím většinou dochází ke ztrátě jejich efektivity.

- **Absence algoritmu na řešení ad-hoc okolních vlivů.**

Stávající algoritmy alokace lidských zdrojů typicky neodrážejí náhodné okolních vlivy v rámci projektů. Nenadálé situace k projektovému řízení patří ve všech metodikách řízení projektů. Za ideální lze považovat projekt bez změnových požadavků s dostatečně vypracovaným registrem rizik, kam projektový manažer, mimo projektová rizika začlení i rizika týkající se interních lidských zdrojů, jako je například nemoc, realokace zdroje na jiný projekt v rámci jeho specializace, realokace s ohledem na místo výkonu práce zdroje nebo jeho bydliště apod. Tyto ad-hoc vlivy mohou změnit projektový plán, kvalitu nebo rozpočet podle vztahu hodnot v troj-imperativu (Mohanty & Malaya, 2011).

- **Podceňování/nerespektování stochastického charakteru činností vykonávaných lidskými zdroji a nevyužívání potenciálu vyplývajícího z jejich variability.**

Stávající algoritmy jsou založeny na statickém modelu plánování lidských zdrojů. Pokud se v průběhu projektu vyskytnou u lidského zdroje nějaké neplánované změny, algoritmus není schopen tyto změny reflektovat, neumí na ně jakýmkoli způsobem upozornit, ani je jinak do systému řízení promítnout. Typickým je tedy přiřazení lidských zdrojů pevně na daný úkol bez jeho změny v průběhu projektu (z pohledu automatizace, vyhodnocování, doporučení v rámci softwarového nástroje).

- **Nedostatečná úroveň vyvažování využívání zdrojů založená na statickém modelu plánování lidských zdrojů.**

Údaje, které poskytují nástroje na podporu projektového řízení, informují projektového nebo resource manažera o vytíženosti nebo volné kapacitě lidského zdroje. Tyto údaje jsou však získávány pouze na základě rozpracovaných projektů a projektů ve stavu plánování, které již mají zarezervované kapacity jednotlivých zdrojů. Právě u těchto projektů by měly být využívány veškeré další relevantní informace sloužící k jejich optimálnímu plánování a vyvažování lidských zdrojů (například prioritizace úkolů, projektů, ale i aktuální schopnosti a zájmy lidských zdrojů v konkrétních oblastech). Je také důležité zmínit, že samotná dostupnost zdroje neznamená, že právě tento zdroj je pro požadovaný úkol v rámci projektu vhodný.

- **Nemožnost uvolňování zdrojů z projektů a jejich následné realokace na projekt s vyšší prioritou a vyšší relevancí pro daný zdroj s ohledem na jeho odbornost a úroveň kvalifikace vzhledem k požadovanému výkonu.**

Koncept řešení problémů alokace lidských zdrojů

Koncepty řešení jednotlivých problémů alokace lidských zdrojů, uvedených v předcházející kapitole, je uveden v tabulce níže.

Tabulka 1 – Koncept řešení problémů alokace lidských zdrojů

Zdroj: autor

Problém	Koncept řešení
Nedostatek informací pro vhodné navržení lidského zdroje pro konkrétní úkol.	Informace o zdrojích by měly být doplňovány z jednotlivých informačních systémů organizace. Tak by byla zajištěna jejich kompletnost a bylo by tak možné efektivně využívat plný potenciál těchto zdrojů, a tím navýšit jejich efektivitu. Informace je možné doplňovat na základě HR dat (například životopis nebo plán osobního rozvoje), dat v adresářové službě organizace (specializace zdrojů, jejich role, nebo například i bydliště) a dalších datových zdrojů.
Lidské zdroje na plný, resp. na částečný (např. poloviční úvazek).	Vhodná alokace zdrojů by měla počítat s neustálým kompenzováním činnosti mezi externími a interními zdroji, a to především s ohledem na dostupné informace, jako jsou zkušenosti, znalosti, zájmy a finanční náklady na zdroj.
Absence algoritmu na řešení ad-hoc okolních vlivů.	Do modelu je možné implementovat zásady a principy systémové dynamiky, a tím zajistit lepší volbu jednotlivých zdrojů k dílčím úkolům v rámci projektu. Alokování zdrojů tak bude probíhat se zpětnou vazbou mezi informační bází s aktuálními hodnotami lidských zdrojů a plánem alokace těchto zdrojů a bude se v rámci monitorovací fáze průběžně přehodnocovat tak, aby nebyla narušena konzistence projektového plánu a vazeb ať už mezi úkoly, tak i mezi projekty.
Podceňování/nerespektování stochastického charakteru činností vykonávaných lidskými zdroji a nevyužívání potenciálu vyplývajícího z jejich variability.	Do evidence zdrojů je možné neustále přidávat aktuální informace o jednotlivých lidských zdrojích (například na základě jejich profesního rozvoje) a s těmito informacemi dále pracovat v rámci variability těchto zdrojů a jejich alokace na projektech v rámci organizace. Tím je možné lépe reagovat na neočekávané vlivy a hrozby v rámci projektů.
Nedostatečná úroveň vyrovnávání založená na statickém modelu plánování lidských zdrojů.	Model vyrovnávání zdrojů v softwarových produktech je založen pouze na informaci o kapacitě a dostupnosti lidského zdroje. Efektivita by mohla být zvýšena pomocí dalších datových zdrojů, například kalendáře zdroje, který je vedený mimo nástroj na podporu projektového řízení a kam se například zadávají dovolené zdroje (které se pak ručně přepisují do nástroje na projektové řízení), dále veškeré události, které mohou vyrovnávání ovlivnit (mimo kalendáře s dovolenými to může být i osobní kalendář zdroje, informace o zdroji – lokalita jeho bydliště a pracoviště apod.).

Problém	Koncept řešení
Nemožnost uvolňování zdrojů z projektů a jejich následné realokace na projekt s vyšší prioritou a vyšší relevancí pro daný zdroj s ohledem na jeho odbornost a úroveň kvalifikace vzhledem k požadovanému výkonu.	Řešením by mohlo být zakomponování vyhodnocovacích algoritmů v rámci běžících projektů v multi-projektovém řízení, kde by výstupem takového algoritmu byly relevantní návrhy na možnou relokaaci zdrojů v rámci nových projektů s vyšší prioritou, změněné priority projektu, zastavení projektu a dalších podobných okolností.

Implementace řešení problémů alokace lidských zdrojů

Implementace řešení problémů alokace lidských zdrojů podle pojetí autora by měla odrážet 2 základní oblasti:

- Informace vztahující se k alokaci lidských zdrojů, které nejsou přímo dostupné v softwaru pro projektové řízení (tj. jsou umístěny v jiných informačních bázích organizace).
- Modifikované algoritmy, které nejsou aktuálně schopné zohlednit více uvedené informace umístěné mimo software řízení projektu.

Jednotlivé problémy by se dali částečně eliminovat integrací jednotlivých informačních systémů v rámci organizace (informační systémy, které mají souvislost s řízením projektů a lidských zdrojů) – tím by se dosáhlo dostatečného množství informací o zdrojích. V rámci propojení jednotlivých informací je možné zvolit produkt uvedený v úvodu tohoto příspěvku (Microsoft Project Server) a propojit ho s dalšími podnikovými systémy. Pro integraci informací lze použít například SharePoint Server nebo Forefront Identity Management, což je HR systém pro správu osob v organizaci.

Zároveň pokládám za vhodné modifikovat algoritmus alokace lidských zdrojů na projekt ze statického na dynamický. Tím budou nastaveny systémové prvky, které budou vyhodnocovat jednotlivé proměnné a budou tyto informace přenášet zpět na začátek (Lyneis & Ford, 2007). Pro model (algoritmus) vycházející ze zpětné vazby při alokaci a řízení lidských zdrojů, přichází v úvahu použití evolučních algoritmů (EA), které jsou poslední dobou často používané pro široká spektra optimalizací. Jedním z nich je například genetický algoritmus (GA) nebo jeho modifikovaná verze, tzv. „multi-objective Genetic algoritmus (moGA)“. Na základě výzkumu Tareka Hegazyho (1999), který se zabýval multikriteriální optimalizací přidělování zdrojů a jejich vyrovňáváním pomocí techniky genetických algoritmů, byl naprogramován skript (programový kód), který lze použít v nástroji Microsoft Project. Tento skript řeší alokaci zdrojů a jejich vyrovňávání. Ani zde však není zohledněna optimalizace alokování lidských zdrojů na základě doplňujících informací o zdrojích a jejich „ad-hoc“ možnostech vyrovňávání. To je možné řešit až na základě serverových technologií, nikoli off-line klientských programů. Ve zmíněném skriptu je dále využívána možnost prioritizace úkolů ze strany uživatele. „Ostatní softwarové systémy, jako například Primavera, většinou toto neumožňují a vyžadují tedy nějakou další manipulaci“⁷ (Hegazy Tarek, 1999; Wen & Lin, 2008).

Závěr

Cílem příspěvku je identifikovat základní problémy projektového managementu v rámci alokace lidských zdrojů, jejich řízení v rámci plánování a exekuce projektu/projektů v multi-projektovém prostředí organizace a navrhnout možný koncept řešení a stručně nastínit implementaci ve stávajících algoritmech. Téměř všechna tato řešení se opírají o myšlenku přepracování modelu alokace a řízení zdrojů, se kterým se v podobě problémů setkává nemalé procento firem na českém i zahraničním trhu. Model neboli algoritmus alokace lidských zdrojů není dostatečně propracován ani v současně využívaných nástrojích na podporu projektového řízení. Lze doporučit implementaci principů systémové dynamiky a dodatečných informačních zdrojů na podporu rozhodování. Na základě tohoto doporučení je podle názoru autora možné očekávat zvýšení efektivity celého plánovacího a vyvažovacího procesu při alokaci zdrojů a při projektovém řízení v multi-projektovém prostředí obecně.

⁷ Pozn. překlad autora.

Literatura

- Bloch, M., Blumberg, S., & Laartz, J. (2012). Delivering large-scale IT projects on time, on budget, and on value. Retrieved from http://www.mckinsey.com/insights/business_technology/delivering_large-scale_it_projects_on_time_on_budget_and_on_value
- Bowen, R. (2012). Dealing With Resource Allocation and Overallocation. Retrieved from <http://www.brighthubpm.com/resource-management/11638-resource-allocation-do-you-overwork-your-resources/>
- Dvořák, D., & Lichtenberg Lisý, D. (2013). Matematické metody v projektovém managementu. *IT SYSTEMS*. Retrieved from <http://www.systemonline.cz/clanky/matematicke-metody-v-projektovem-managementu.htm>
- Hart, S. (2012). Using MS Project Server for Resource Management | PM-Foundations on WordPress.com. Retrieved from <http://pm-foundations.com/2012/11/01/using-ms-project-server-for-resource-management/>
- Hegazy Tarek. (1999). Optimization of resource allocation and leveling using genetic algorithms. *JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT*, MAY/JUNE 1(June), 167–175.
- Lyneis, J., & Ford, D. (2007). System dynamics applied to project management: a survey, assessment, and directions for future research. *System Dynamics Review*, 23(2), 157–189. doi:10.1002/sdr
- Mohanty, S., & Malaya, K. N. (2011). Optimization Model in Human Resource Management for Job Allocation in ICT Project. *International Journal of the Computer, the Internet and Management*, 19(3), 21–27.
- Raunak, M., & Osterweil, L. (2005). Effective resource allocation for process simulation: A position paper. ... *Workshop on Software Process Simulation ...*, 1–4. Retrieved from [http://www.umass.edu/eei/EEI Website Articles/Effective Resource Allocation for Process Simulation- A Position Paper.pdf](http://www.umass.edu/eei/EEI%20Website%20Articles/Effective%20Resource%20Allocation%20for%20Process%20Simulation-%20A%20Position%20Paper.pdf)
- Roberto, M.-D., Bruno, B., & Jose-Luis, F. (1976). Computer Aided Systems Theory – EUROCAST 2001. *A Selection of Papers from the 8th International Workshop on Computer Aided Systems Theory*, 8, 679. Retrieved from <http://www.ulb.tu-darmstadt.de/tocs/79304567.pdf>
- Wen, F., & Lin, C.-M. (2008). Multistage Human Resource Allocation for Software Development by Multiobjective Genetic Algorithm. *The Open Applied Mathematics Journal*, 2(1), 95–103. doi:10.2174/1874114200802010095

JEL Classification: M15, O22

Summary

Human resource allocations in project management

At present, most companies manage projects through various support tools. A large proportion of those companies use different already proven mechanisms to create time estimates and allocation of human resources to tasks. However none of software tools is sophisticated enough to solve one of the most common problem of our time - the human resource allocation - at an advanced level.

The work is devoted to the standard methods of human resource allocation and allocation with respect to other ad-hoc influences and this issue is transferred to the environment software tools to project management support. It deals with the main shortcomings of the allocation process, human resources and potential impacts on the outcome of one or more concurrent projects.

To address the current gaps author recommends the use of other data sources, which would have supplied the necessary information into the planning process (or in a software tool to support project management) and follow-up treatment algorithm for the allocation and balancing human resources.

Key words: Enterprise Project Management, Resource Management, PERT, Monte Carlo, IS Integration.

Agile Outsourcing of Software Development

Aziz Ahmad Rais

aziz.rais@vse.cz

Ph.D. Students of Informatics

Supervisor: Ing. Rudolf Pecinovský, CSc, (rudolf.pecinovsky@i.cz)

Abstract: Outsourcing is a known strategy in the business informatics. Outsourcing is still in one of the IT trends therefore still many companies outsource besides others their software development too. It is also known that there are connected some risk with software development, when outsourcing a software development these risks become more critical.

Outsourcing software development's risks mitigation is difficult comparing to an internal software development. Because in outsourcing of software development plays role at least two parties (customer and vendor) that are bound by a contract to cooperate. This means that all parties involved in outsourcing of software development have to protect their own interests. If there is not proper software development methodology used during outsourcing of software development to prevent from the risks and have possibilities to identify and eliminate the risks on time, then it will be very difficult to solve the problems caused by the risks without additional investments.

The goal this article is to introduce a new agile software development methodology that can better identify and eliminate these risks in both internal as well as in outsourcing of software development.

Key words: Outsourcing of software development, agile software development methodology, and interface based development.

Introduction

Software development could be outsourced partially or fully to one or more vendors. In this article is considered full outsourcing of software development. A software development will be understood a software development life cycle as defined by CMMI for dev. see the details in [Buchalceková, A. 2009].

Thus agile outsourcing of software development will be understood an outsourced software development that uses an agile software development methodology.

In order to understand the need for new agile software development methodology, it is important to understand the risks connected to outsourcing of software development. Software development could be outsourced to one or more vendors. The goal of this article is to not solve the how many parties are involved in the software development outsourcing, but only outsource the software development to one vendor.

When outsourcing a software development we can get very similar software development as internal. Here from we can conclude that most of the risks that are connected to internal software development will appear also in outsourcing of software development. This implication is could be verified by comparing risks identified by CHAOS report of Sandish Group and risks connected to outsourcing by [Wan, J., et al, 2010].

The following risks of IT projects are identified by CHAOS report of Sandish Group. Although the Sandish Group report doesn't specify whether the IT projects are outsourcing projects or in-house, but the case study provided by the report implies that all the risks below are related to the in-house projects:

- Incomplete Requirements (Unclear Objectives)
- Changing Requirements & Specifications
- Unrealistic Expectations
- User Involvement 15.9%
- Technology competence
- Lack of Resources
- Lack of Executive Support 13.9%
- Lack of Planning (e.g. Unrealistic Time Frames)

- Lack of IT Management
- Other

Here is a list of risks identified with IT projects outsourcing by [Wan, J., et al, 2010]:

- The products requirement are unclear
- The requirement changed continually
- Software maintenance and technical support
- Personnel capacity, resources
- Lack of responsibility
- Service quality
- Software test plans

To compare risks of both types of software development we can see risks connected to requirements (unclear requirements, frequently changing requirements, and unclear objectives). After analyzing the rest of risks in detail, we can conclude that the remaining risk in both types of development has to do with technical quality of software (e.g. Technology competence, Software maintenance and technical support) and quality of software development methodology.

From the CHAOS report of Sandish Group implies also that mostly project failure is cause by requirements risks. This means that requirements risks could be root cause of failure and also can affect technical quality. Thus the main goal of the new methodology introduced in the next chapter is to provide a solution that can better mitigate the business and technical risks.

Interface based development

The goal of the methodology is to simplify analysis and design of software development outsourcing and unifying views: design, architecture design, implementation, and testing by using interfaces. The interface will be understood as follow: The interface of a given entity (= of a program's part – a module, a class, a method...) specifies what the given entity knows and how to communicate with it [Pecinovský Rudolf 2013]. Interfaces can also unify the views of: the customer's stakeholders, the customer's quality assurance, the vendor's quality assurance, and the vendor's technical team (architect, developer).

Here we should emphasize that we speak about a general interface specifying a general properties and functionality of program entities (components, classes, objects, methods, attributes ...). The concrete realization of the designed interfaces depends on the programming language and platform in use.

One of the main advantages of IBD is reducing the redundant analysis and design work. Starting development with the identification of interfaces will remove the need for use cases. It will force development team to think in functionality by creating interfaces and not thinking about how to implement a specific functionality. As interface specifies an abstract data type, there is no need for further identification of conceptual classes as it happen in a classical OOA&D. To better understand this advantage there is a need to describe in short how object-oriented analysis and design (OOA&D) works for details description of OOA&D see [McLaughlin, B. et al.]. Object-oriented analysis process inputs are a requirements list or a features list, sometimes with business process or wireframes; then the following steps are needed to complete OOA&D:

- 1) Description of use cases and drawing use case diagrams (step-1).
- 2) Breaking up the functionalities or use cases in modules (step-2).
- 3) Domain analysis, identify business domain and domain objects (step-3).
- 4) Primary design (step-4).
 - a. Identification of list of conceptual components and entities based on identifying nouns and verbs.
 - b. Transformation of the conceptual components and entities into classes.
- 5) Architecture: organize components to identify the relation between them (step-5).

The inputs to the IBD will be the same: requirements list or features list, wireframes and business process models. With interface based development (IBD) the following steps need to be taken:

- 1) Identification of interfaces and drawing sequence diagrams based on interfaces (step-1).

- 2) No step 2 (step-2 in classical OOA&D) with additional care about modularity is needed because the defined interfaces already guarantee the modularity. (The organization of the interfaces into packages is better to do as part of architecture design.).
- 3) Identification of domain objects based on interfaces. The domain objects are the inputs of the interface (step-3). This is also not needed to do separately. It is easier to identify domain objects from interface than from a use case.
- 4) No design step (step-4 in classical OOA&D) is needed, as the interfaces already represent abstract data types.
- 5) Architecture: Organize the interfaces into layers to identify the relation between them (step-5).

From the above process it is evident that the interface base development helps in reducing three steps (2, 3 and 4) from the classical OOA&D process.

Example to compare OOA&D with IBD

To prove the above and figure out why there is no need for use cases, let's take a look at the example that is provided in the [McLaughlin, B. et al.].

Features list

The features list of the Route Finder application:

- 1) We have to be able to represent a subway line, and the stations along that line
- 2) We must be able to load multiple subway lines into the program, including overlapping lines.
- 3) We need to be able to figure out a valid path between any two stations on any line.
- 4) We need to be able to print out a route between two stations as a set of directions.

OOA&D using use cases

This chapter and its sub-chapters will describe in short the OOA&D described in the book [McLaughlin, B. et al.].

Use case diagram (step-1)

The above four features see “features list” chapter could be described in two use cases.

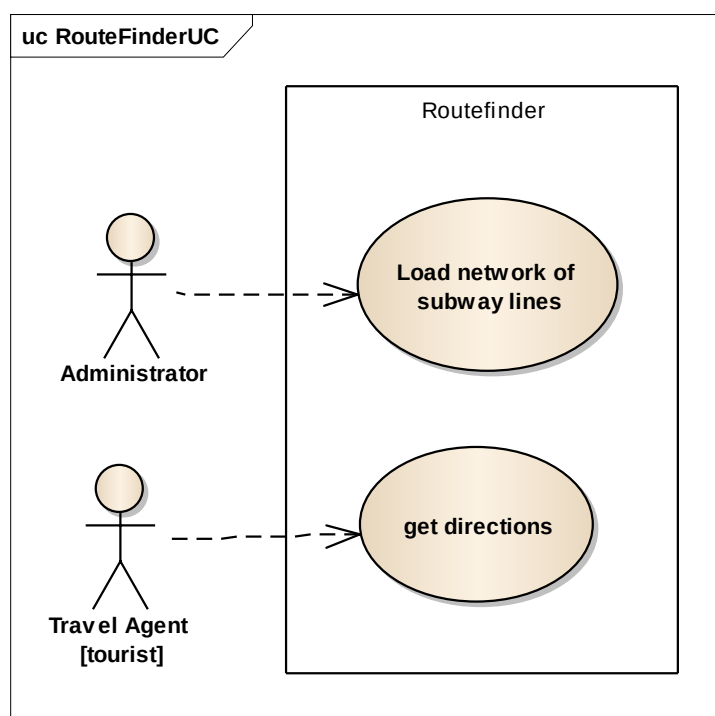


Diagram 7-1: use case diagram of application Route finder

Source: author

- 1) The scenario of use case “Load network of subway lines”:
- 2) The administrator supplies a file of station and lines.
- 3) The system reads in the name of the station.
- 4) The system validates that the station doesn't already exist.
- 5) The system adds the new station to the subway.
- 6) The system repeats steps 2-4 until all the stations are added.
- 7) The system reads in the name of a line to add.
- 8) The system reads in two stations that are connected.
- 9) The system validates that the stations exist.
- 10) The system creates a new connection between the two stations, going in both directions on the current line.
- 11) The system repeats steps 7-9 until the lines are completed.
- 12) The system repeats steps 6-10 until the lines are entered.

The scenario of use case “get directions”

- 1) The travel agent gives the system a starting station and station to travel to.
- 2) The system validates that the starting and ending station both exist on the subway.
- 3) The system calculates a route from the starting station to the ending station.
- 4) The system prints out the route it calculated.

Break up the functionalities or use cases in modules (step-2)

There are three packages (modules) identified:

- 1) Loader: this package contains loading subway data from file.
- 2) Printer: printing the search result, so printing objects.
- 3) Subway: domain objects.
- 4) Test: unit tests.

Domain analysis, identify business domain and domain objects (step-3)

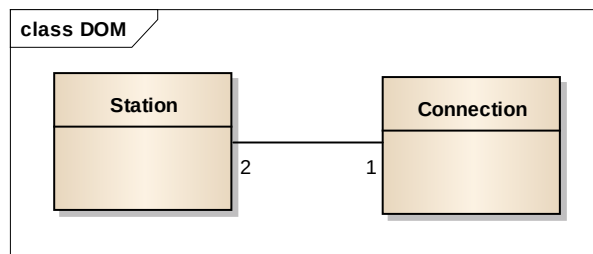


Diagram 2: domain object model of application Route finder

Source: author

Primary design (step-4)

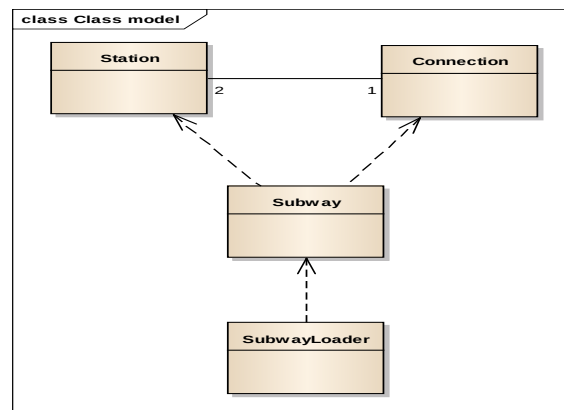


Diagram 3: class diagram of application Route finder

Source: author

OOA&D in interface base development

Features 1 and 2 imply that the application Route Finder have to provide an interface for loading the subway information (station and line) into the application. Feature 3 and 4 imply that the application Route Finder has to provide interfaces for searching a route between two stations and print the result. Further user types would be needed to authorize the user when interacting with the application. The goal of authorization is to prevent that – every user type cannot load any type of subway information into the system.

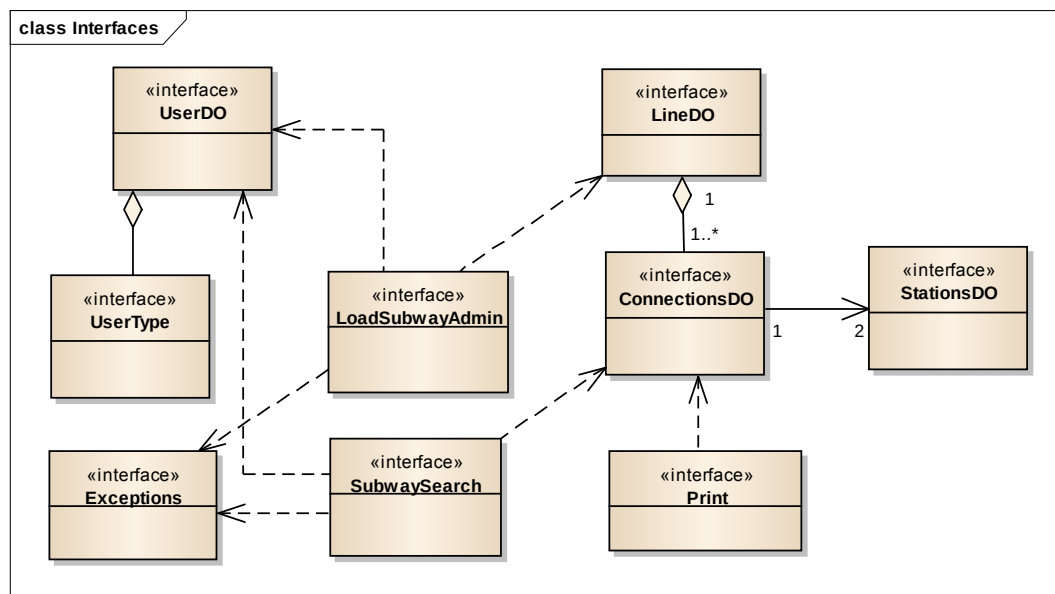


Diagram 4: interface diagram of application Route finder

Source: author

Comparing the interface diagram with use case analysis we can see that a use cases can be represented by interfaces and domain objects are obtained by sending messages to that interfaces.

As the requirements list is short and doesn't describe everything, the goal of IBD is to complete the business requirements with identification of additional interface. Additional interfaces could be identifying when trying to capture how the user will interact with Route Finder application. In the analysis of business process "Load network of subway lines" by drawing a sequence diagram we can see the need for additional interfaces. Instance of each interface will represent an object in sequence diagram.

This way we can complete the business analysis and the same time thinking about what further a software application needs. If we miss anything in the requirements list this way we can easily identify the missing requirements and involve customer to decide and provide additional information. Notice also that with sequence diagrams we have already identified the process view of the application architecture.

The sequence diagram representing scenario of use case "Load network of subway lines":

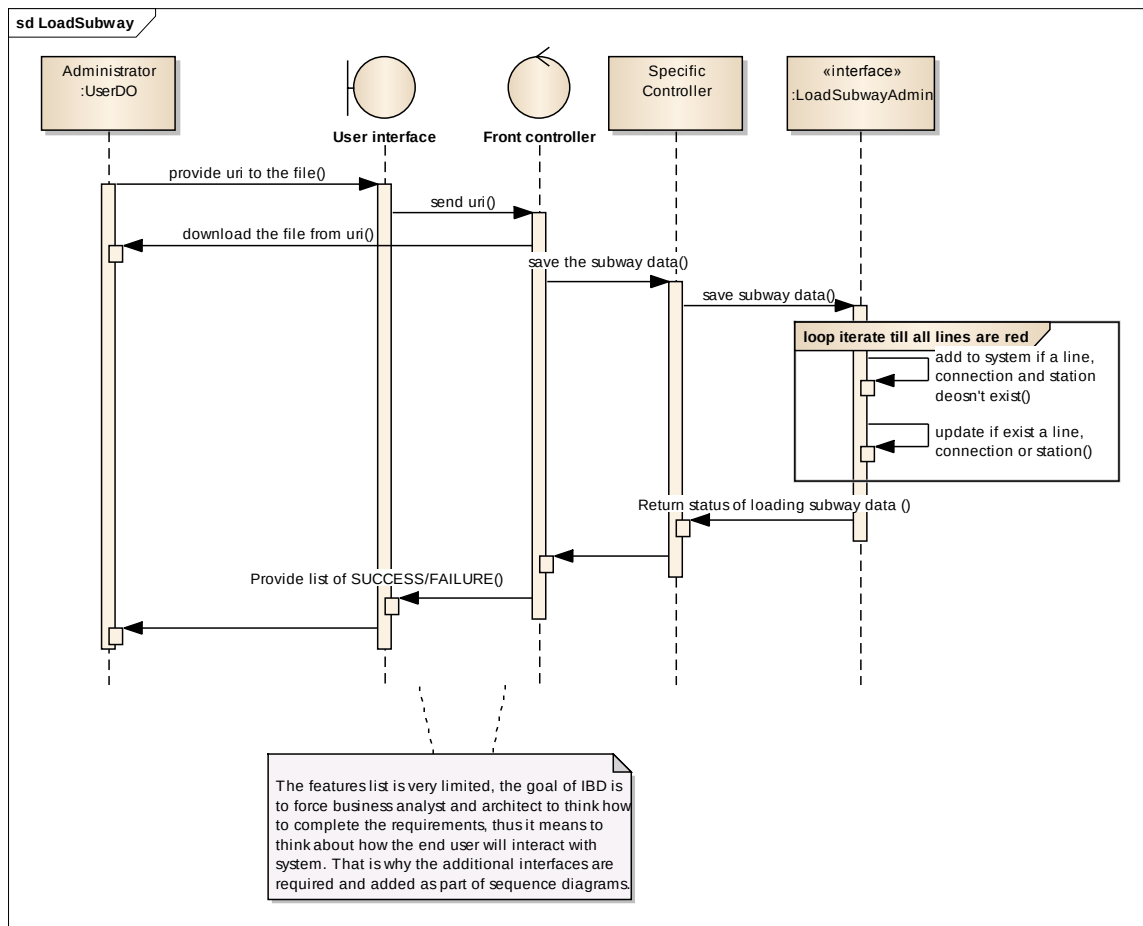


Diagram 5: Scenario one of application Route finder

Source: author

Because we are in the phase of analysis of requirements, we can add any additional objects into the sequence diagram. But any of these additional objects has to be represented by some interface, framework or third party component. So the sequence diagrams in the chapter use view and controller from MVC design pattern that will be provided by some framework which will be covered in technology architecture. Technology architecture is one of the views of application architecture that will be covered by architecture method that will be provided by dissertation.

The “search a connection and print” scenario:

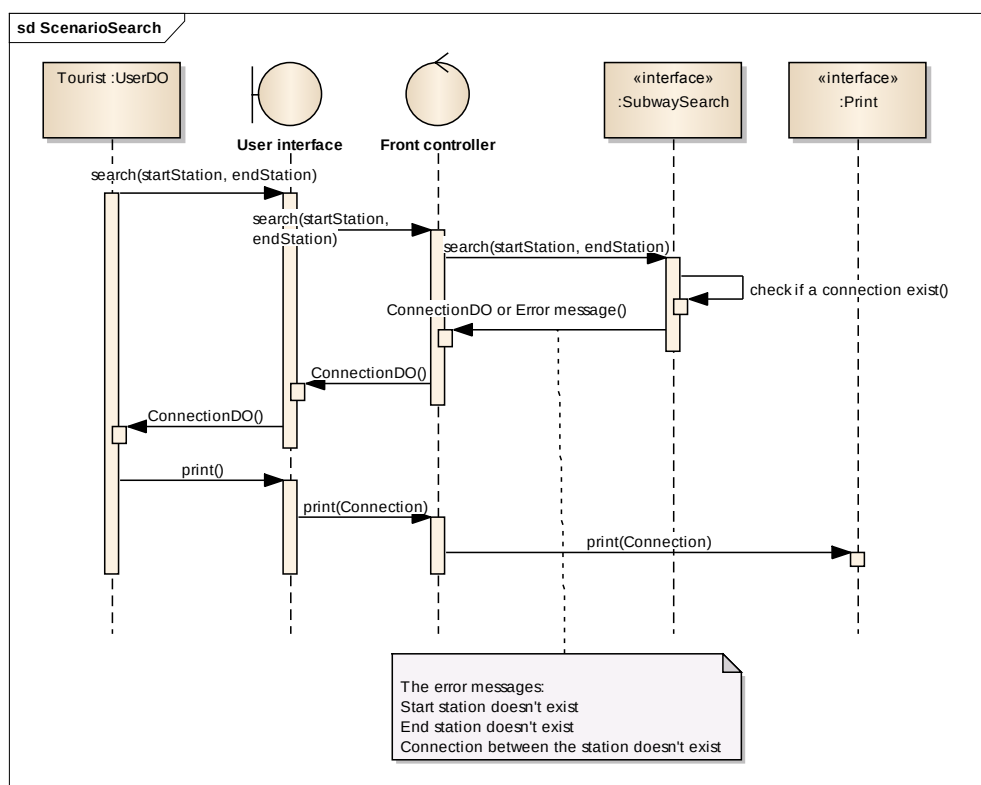


Diagram 6: Scenario two of application Route finder

Source: author

Notice that in the sequence diagram 7-6 there is missing a specific controller as in the first business process. It is left because to demonstrate that in the architecture design could be specified additional interfaces to meet some additional quality requirements. But adding any additional interface should not affect the process of searching a connection between two stations.

Break up the functionalities or use cases in modules (step-2)

The interfaces identified above could be divided into the following packages. Dividing interfaces into packages in IBD will be done as part of architecture but it is shown here only for comparison with OOA in chapter [OOA&D using use cases]:

- 1) Loader: this package contains loading subway data from file and searching connections.
- 2) Printer: printing the search result, so printing objects.
- 3) Domain objects: domain objects.
- 4) Exceptions.
- 5) Test: unit tests.

Domain analysis, identify business domain and domain objects (step-3)

Domain objects are identified already as part of interfaces- thus no need for further diagrams.

Primary design (step-4)

The identification of components is easy because, a component is a replaceable, executable piece of a larger system whose implementation details are hidden. The functionality provided by a component is specified by a set of provided interfaces that the component realizes (see "Black-Box View"). In addition to providing interfaces, a component may require interfaces in order to function. These are called required interfaces [Pilone Dan, Pitman Neil]. Thus there is no need for class diagrams, component diagram will be important only if application software has a lot of interfaces and there is required a component view of whole system. In such we can easily identify component with provided interfaces that are already identified in analysis.

Architecture: Organize the interfaces into layers (step-5)

The goal of this chapter is to show the consistency of requirements and architecture design in IBD.

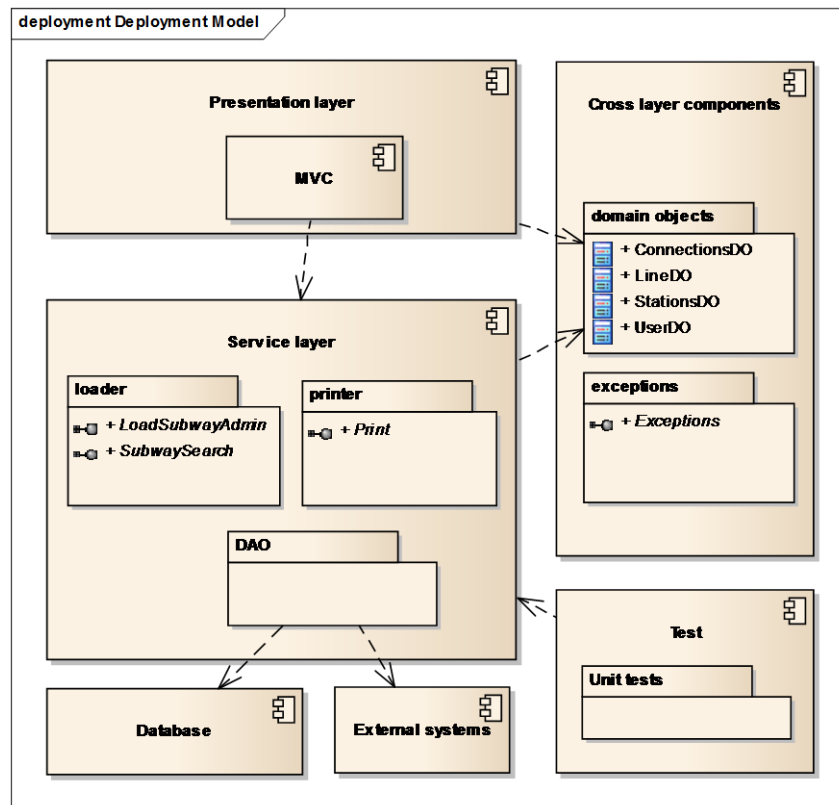


Diagram 7: The application architecture of Route finder application

Source: author

Different views of interface

Interfaces are easy to organize in packages and in layers. For quality assurance purposes the interfaces are easy to use for unit test generation, integration test. Interfaces increase the testability of the system. Interfaces are core concept of object-oriented programming and most developers are familiar with this concept. Therefore it is easy for developers to implement them in a concrete class or classes.

The conclusion is that interfaces unify the view of developers, architects, testers and designers.

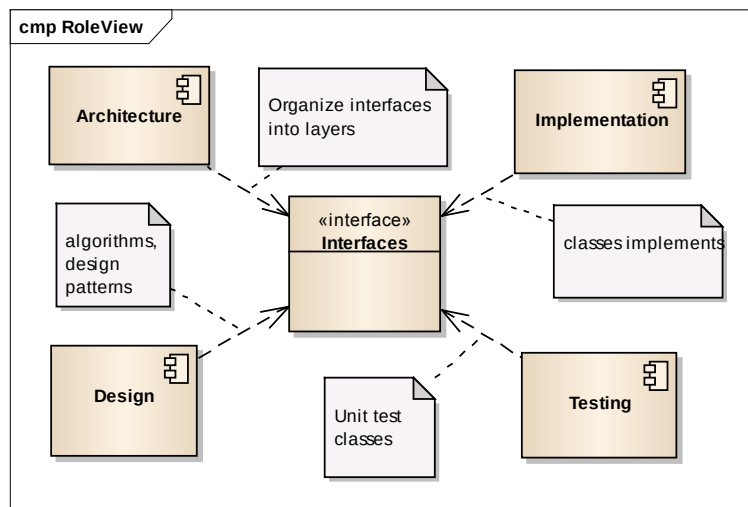


Diagram 8: describes the consistency between: architecture, design, implementation and testing

Source: author

Conclusion

In the traditional OOA&D methodology first all the requirements are done by identifying the use cases, thus comparing the phase with how requirements can be analyzed in IBD we can see that in IBD there are analyzed not only the requirements but also interfaces of all the main business components. These interfaces also help in identifying domain objects. But with traditional OOA&D the main components and domain objects are identified in the next phases of software development. It is generally known that the use cases are written in the natural languages, which has limitation from application software development point of view. The limitation is that the natural languages cannot describe the application software always as unique and unambiguous. This causes misunderstanding between business and software developers.

Another limitation of the use cases is the transformation of use cases into software component. This transformation causes inconsistency between requirements and software functionality. This transformation also causes that always business modularity cannot be followed by software modularity. From the limitation of use cases implies that the application software is always analyzed twice in two different languages (natural and object oriented programming language e.g. Java, C#...). That is not only a root cause of misunderstanding but also a redundant work.

That is why IBD suggest describing application software in the programming language feature that describes functionality of components of the application software and that it is core feature to the object oriented programming language.

With the use of interface IBD solves not only requirements risks but also it provides mechanism to keep in consistency the requirements and application software design. With IBD it is easy to track the changes in the requirements and support model driven architecture and development principles too.

The IBD is suitable not only for outsourcing of software development but also internal software development. This implies from that the IBD solves risks that are similar in both types of software development.

References

- Manifesto for Agile Software Development* [Online], 2001. available on: <http://agilemanifesto.org/>
- BUCHALCEVOVA Alena. *Methodologies of information systems development*. Oeconomica – Prague 2009. ISBN: 978-80-245-1540-3
- WAN Jiangping, WAN Dan, ZHANG Hui: *Case Study on Business Risk Management for Software Outsourcing Service Provider with ISM: Technology and Investment*, 2010, 1, 257-266 doi:10.4236/ti.2010.14033 Published Online November 2010 (<http://www.SciRP.org/journal/ti>)
- PECINOVSKY Rudolf: *OOP – Learn Object Oriented Thinking and Programming*, 2013. Eva & Tomas Bruckner publishing. ISBN 978-80-904661-9-7
- MCLAUGHLIN Brett, POLLICE Gary, WEST David. *Head First Object-Oriented Analysis and Design*. O'Reilly Media ISBN: 978-0-596-00867-3
- PILONE Dan, PITMAN Neil, *UML 2.0 in a Nutshell*. O'Reilly. June 2005, ISBN: 0-596-00795-7

Přispívá internet k rozvoji tržního systému?

Václav Řezníček

vaclav.reznicek@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: doc. Ing. Stanislava Mildeová, CSc., (mildeova@vse.cz)

Abstrakt: Tento příspěvek, jak samotný jeho název napovídá, pokládá otázku, zda internet jakožto významný fenomén informatizace (globální informační systém) přispívá k rozvoji tržního systému. Výzkum je postaven na aplikaci teorie rakouské ekonomické školy, jež je založena na metodologickém apriorismu, metodologickém dualismu a individualismu. Tato škola chápe ekonomiku jako komplexní (spontánní) systém, který na rozdíl od člověkem účelově vytvářených artefaktů (organizací) není produktem záměrné individuální činnosti, ale vzniká jako výsledek nikým (shora) neplánované egoistické činnosti milionů individuí, které se podřizují jistým abstraktně-formálním pravidlům. V souladu s kybernetikou druhého řádu pohlíží na konkrétního jednotlivce jako na prvek, který nestojí mimo systém, ale je jeho aktivní součástí, a jakožto prvek systému komplexního trpí ex definitione omezenou znalostí faktů.

Uvedená teoretická perspektiva poskytuje unikátní pohled na danou problematiku, zejména pokud jde o přístup jednotlivců k informacím v dnešních, vlivem internetu změněných podmínkách. V textu je ukázáno, že internet umožňuje nejen díky snazší komparabilitě cen efektivnější rozhodování v rámci jinak složitějšího a v detailech neuchopitelného systému, ale rovněž může být i důležitým nástrojem napomáhajícím vynucovat dodržování zmíněných „pravidel chování“. Na druhé straně je v příspěvku diskutován tržní, zájmově podmíněný charakter poskytování informací, který má potenciál resultovat v rozšiřování informací zkreslených a mylných. V daném kontextu je pak poukázáno na problematiku vzdělávání a na roli lidské znalosti, neboť právě ta se zdá být co do budoucích prognóz určující.

Klíčová slova: Informace, znalost, internet, tržní systém, komplexní systém, spontánní řád, ekonomika.

Úvod

Předkládaný text se zabývá otázkou, zda internet jakožto významný fenomén informatizace (ve smyslu pronikání informačních technologií do ve společnosti probíhajících procesů) přispívá k rozvoji tržního systému. Jeho cílem není vyřčení jednoduché odpovědi „ano“ či „ne“, ale snaží se ukázat, jak je vůbec možné k takové (v literatuře prakticky neřešené) otázce seriózně přistoupit. Jelikož se jedná o prezentaci dílčích výsledků disertační práce autora, nevyhne se tento příspěvek určitému „vytržení z kontextu“, neboť ve stanoveném formátu není možné, aby byl vyčerpávající. S tím souvisí skutečnost, že mnohé pouze načrtává a odvolává se přitom na autorem již publikované texty, případně na další výzkum, který bude na základě provedené reflexe a analýzy realizován a následně představen ve finální podobě disertační práce.

Struktura příspěvku, který byl nazván otázkou „Přispívá internet k rozvoji tržního systému?“, je následující: Nejprve jsou představena relevantní metodologická a teoretická východiska. Tuto část tvoří zejména vymezení základních fundamentů teorie rakouské ekonomické školy a vydefinování užívaných pojmů. Těmi základními jsou samozřejmě tržní systém a internet. Dále je prostor věnován stěžejní části tohoto textu, tedy aplikaci internetu v rámci tržních procesů a jeho vlivu na rozvoj tržního systému v prezentované teoretické perspektivě. Nakonec jsou naznačeny příspěvek přesahující souvislosti směrem k problematice lidské znalosti a otázce vzdělávání.

Autor této statě zde danou ekonomickou teorii, která byla formulována roky či spíše desetiletí před nástupem internetu, nehodnotí a nesrovnává s jinými přístupy (respektive dojde-li ke srovnání, je toto prostředkem a nikoli cílem autora),¹ ale snaží se ji aplikovat v dnešních změněných podmínkách, stejně jako přispět k rozvoji směrem k hlubší diskusi nad danou problematikou a poskytnout závěry pro současnou praxi.

¹ Teorie rakouské školy však nebyla zvolena náhodně a v disertační práci autora tohoto textu je věnován patřičný prostor otázce této volby a jejímu vysvětlení, kde je ukázáno, že přístupy rakouských teoretiků ve srovnání s objektivistickým neoklasickým paradigmatem lépe korespondují s autorem prezentovaným pojetím informace a individuální lidské znalosti.

Metodologická a teoretická východiska

Předkládaný text vychází především z analýzy primární literatury, tedy z díla významného představitele rakouské ekonomické školy (dále jen rakouské školy), Friedricha Augusta von Hayeka, jako je (Hayek, 1945), (Hayek, 1955), (Hayek, 1960), (Hayek, 1995), (Hayek, 1998), (Hayek, 2008) či (Hayek, 2011) a z díla Ludwiga von Misesa (1969) a (2007). Jeho sepsání předcházelo několikaleté studium těchto a souvisejících prací. Mimo uvedené primární prameny se pak opírá rovněž o analýzy Hayekových myšlenek sepsané Janem Pavlíkem (2004a) a (2004b) nebo o příspěvky několika autorů zveřejněné v knize (Schwarz za kol., 2000). Sám autor této práce již publikoval několik dané problematiky se (do)týkajících výstupů. V článku publikovaném v roce 2011 (Řezníček, 2011) diskutoval zásadní ideu Hayekova díla, totiž ideu omezenosti našeho poznání. Jedna z dalších publikací (Řezníček, Smutný, 2012b) se pak věnovala otázce složitosti systémů v jejich (pro rakouskou školu významné) distinkci na systémy přirozené a umělé.

Za základní metodologické východisko rakouské školy lze označit apriorismus. Jde o přístup k ekonomii nikoli jako k vědě empirické, ale jako k vědě apriorní. „Ježto apriorní ekonomie odhaluje podmínky možnosti ekonomické zkušenosti (jež jsou eo ipso neempirickými podmínkami, neboť to, co umožňuje zkušenost, nemůže samo být zkušeností), nejsou její jednotlivé teoremy přístupné žádnému způsobu empirické verifikace či falzifikace“ (Pavlík, 2002, str. 12). Tato skutečnost nezřídka kdy stačí k tomu, aby mnozí vědci ekonomové metodologický apriorismus zavrhnuli s tím, že se jedná o nevědecký přístup. Ve spojení s metodologickým dualismem kladoucím důraz na principiální odlišnost metod² používaných v přírodních a ve společenských vědách (Hayek, 1995; Mises, 1969; Mises, 2007) zároveň dochází u představitelů rakouské či neorakouské školy k radikálnímu zamítnutí kvantifikace a matematizace jak premis, tak implikací ekonomických teorií. Zmíněné, stejně jako význam přisuzovaný aprioristům jedinečnosti a neopakovatelnosti historických událostí, implikuje, že se ekonomická věda v daném pojetí zříká predikční funkce vědeckého poznání. Jak poznamenává Jan Pavlík (2002, str. 13), absence predikční schopnosti ekonomie „musí nevyhnutelně iritovat standardního neoklasického ekonoma, který se v souladu se svým pozitivistickým³ školením domnívá, že tvrzení ekonomie jsou pouhé induktivní hypotézy a že jeho úkol spočívá ve výstavbě testovatelných matematických modelů, mezi nimiž probíhá selekce na základě jejich relativní predikční síly; pokud by měl akceptovat závěry ekonomického apriorismu, musel by si přiznat, že veškerá jeho profesní aktivita je nesmyslná a zbytečná – a musel by také opustit mentalitu sociálního inženýrství, která je pozitivistickému postoji inherentní.“

V souladu s kybernetikou druhého řádu (Heylingen, Joslyn, 2004) klade rakouská škola důraz na roli pozorovatele, který je sám součástí daného systému a rozhoduje se o svém jednání. Hayek dospěl k tvrzení, že individuum trpí osudovým zakletím do parciality,⁴ přičemž zároveň touží saturovat potřebu porozumět celku. Vzniklý rozpor se pak člověku daří překonávat namísto bezvýsledné snahy o detailní porozumění jednotlivostem především tak, že se spoléhá na schopnost abstrakce, tedy na schopnost pochopit obecné principy a pravidla. Tomáš Ježek (Schwarz za kol., 2000, str. 130) k tomu poznamenává, že „lidské poznání se zdokonalovalo nikoli hlavně tím, že bylo s to pojmut stále větší množství faktů, nýbrž tím, že se spoléhalo na abstrakci jako na nepostradatelný prostředek myslí, jenž jí dává schopnost zacházet s realitou, kterou nemůže plně ovládnout.“ Hayekovo tázání se po způsobech poznávání je zřetelně instrumentální povahy a to sice ve smyslu, že je přípravou k tomu, aby mohl podat důkaz o nemožnosti vědeckého centrálního celospolečenského řízení (Řezníček, 2011). Komplexní charakter abstraktního (endogenního) tržního řádu⁵ (kosmos), který budeme v této práci nazývat tržním systémem, je u Hayeka v opozici proti (exogennímu) institucionálnímu řádu organizace (taxis). Organizace je vědomě a účelově (záměrně) vytvořený systém (lidský artefakt), kterému jsme schopni, co by jeho tvůrci, v detailech porozumět a můžeme jej účinně řídit. Trh je naproti tomu spontánní řád, který je komplexním

² Hayek (1955) ukazuje, že různé vědy mají různou „vysvětlovací“ schopnost.

³ Jak uvádí Mises (1969), vědy o lidském jednání (jako je věda ekonomická) musejí odmítnout nikoli determinismus, ale jeho pozitivistické zkreslení (viz Misesův metodologický individualismus).

⁴ Uvedené označení používá Tomáš Ježek ve svém článku publikovaném v knize (Schwarz za kol., 2000, str. 130).

⁵ Hayek (1998, str. 43) vymezuje řád obecně jako „ten stav věcí, v němž se velký počet prvků různých druhů má k sobě navzájem tak, že znalost nějaké prostorové nebo časové části celku nám umožňuje vytvářet správná očekávání týkající se zbytku, nebo alespoň očekávání, která mají dobrou pravděpodobnost, že se ukáží jako správná.“

systémem, jenž má ex definitione v potenci nabýt jakéhokoli stupně složitosti⁶ a který principiálně nelze vysvětlit v detailech. To má za následek problémy při pokusech o účelovou (racionální) konstrukci. Spontánní řád trhu je, jak Hayek (1998) zdůrazňuje, nikým záměrně nevytvořenou abstraktní strukturou založenou na abstraktních vztazích, jež může přetrvávat kvantitativní (i kvalitativní) změnu jejích prvků a která vzniká a funguje díky respektování a podřizování se těchto prvků (jednotlivců a organizací) jistým pravidlům (chování), aniž by tyto přitom znaly konkrétní účelnost dodržování pravidel pro fungování tržního řádu (systému) jako celku. Miliony jednotlivců tak vytvářejí spontánní řád nevědomou interakcí a kombinací nezáměrných důsledků ekonomických činností účelově vedených pouze snahou maximalizovat užitek pro sebe sama. Ján Pavlík (Schwarz za kol., 2000, str. 60-61) k tomu uvádí, že „spontánní řád trhu, který lze označit za systém nevědomého a nezáměrného vzájemného donucování všech všemi k maximalizaci úsilí o efektivní uspokojování potřeb, se projevuje jako předurčená harmonie, tj. v podobě nikým neplánované synchronické a diachronické koordinace nesčíslného množství svobodných lidských aktivit (záměrných pouze k maximalizaci vlastního prospěchu).“ Jedná-li individuum egoisticky v souladu s danými abstraktně-formálními pravidly, získává jeho jednání eticko-právní legitimitu, takže se nemusí starat o ty důsledky svých vlastních činů, které přesahují jeho individuální užitek.

Jak již bylo naznačeno, Hayekova teorie byla formulována pod záměrem dokázat, že centrální ekonomické řízení a plánování není možné a dlouhodobě pro danou ekonomiku představuje zásadní negativum uvrhující ji do nerozvinutosti a primitivismu. Ve své článku o Využití znalostí ve společnosti (Hayek, 1945) argumentuje koncepcí ve společnosti rozptýlených znalostí, kdy ukazuje, že ekonomický problém společnosti nespočívá pouze v otázce alokace disponibilních zdrojů ale zejména v otázce, jak zajistit nejlepší (nejefektivnější) použití (využití) daných zdrojů známých jednotlivým členům dané společnosti pro účely, jejichž relativní význam znají opět pouze tyto jednotlivci. Jedná se o problém využití znalostí, které nejsou nikomu dány ve své úplnosti, neboť neexistuje jednatel, který by měl takové znalosti, aby byl schopen předpovídat jednání a potřeby všech lidí ve společnosti (poznamenejme, že by to zároveň znamenalo předvídat objevy, které ještě nebyly učiněny).⁷ Hayek navíc ve svém psychologickém díle *The Sensory Order* (Hayek, 2011)⁸ podává exaktní důkaz, že lidský mozek (co by komplexní systém) nedokáže vysvětlit své vlastní fungování, a z toho implikuje, že právě proto, že mozek nemůže poznat a předvídat chování (fungování) sebe sama, neboť by se musel jaksi „nadřadit“, není možné, aby jakýkoliv centrální plánovač znal do detailu činnosti všech jednotlivců ve společnosti. I když připouští možnost existence stroje, který by měl vyšší informační kapacitu, než jakou má lidský mozek, tak by takový stroj musel být napojen na mozky všech lidí ve společnosti, aby z nich v reálném čase čerpal informace, což je úvaha jevící se co do své realizace (alespoň prozatím) jako velmi nepravděpodobná.

Jakým způsobem se tedy mají jednotlivé subjekty na trhu (úspěšně) orientovat? Pokud vyjdeme z uvedeného tvrzení, že každé individuum jedná pouze v horizontu svých vlastních potřeb a v zájmu maximalizace svého užtku, pak pro orientaci tržního subjektu v rámci komplexního tržního systému podle Hayeka (2008) poslouží systém cen. Ty informují prvky celého systému o tom, co mají nakupovat, vyrábět atd., tedy dávají impulsy k jejich účelovému jednání. Nezávislý cenový systém je schopen koordinace činnosti milionu jednotlivců, přičemž žádný z účastníků nemusí mít o cenách všechny informace, ale stačí mu určité penzum o skutečnostech (cenách) týkajících se bezprostředně jeho individuálních zájmů. „Cenový systém v konkurenčním prostředí umožňuje podnikatelům, aby sledováním relativně malého množství cen přizpůsobovali svou činnost aktivitám ostatních“ (Hayek, 2008, str. 57). Čím je celek složitější, tím je závislejší na transmisi informací (Hayek, 2008) mezi jednotlivci, jejichž vlastní aktivity jsou koordinovány právě neosobním mechanismem přenášení daných informací (o cenách). Zároveň připomeňme Hayekův důraz na nutnost dodržování pravidel (při vykonávání svých individuálních

⁶ Ján Pavlík (2004a, str. 19) k tomu poznamenává, že dle Hayeka spontánní řády „nemusejí být nutně složité, ale mohou dosáhnout jakéhokoliv stupně složitosti, neboť stupeň složitosti spontánního řádu jakožto nezáměrného produktu lidské činnosti není omezen tím, co může lidská mysl zvládnout.“

⁷ Hayek (1998) upozorňuje, že konkurence je způsob objevování. Trh je tedy procesem objevování. A k výše uvedenému ve své knize *Kontrarevoluce vědy* (Hayek, 1995) dodává, že mnohé ze sporů týkajících se ekonomické teorie i ekonomické politiky mají svůj společný původ v chybném pojetí povahy ekonomického problému společnosti. Toto chybné pojetí samo má pak svůj prapočátek ve zcela nešťastném přenášení myšlenkových návyků, které jsme si vytvořili při našem styku s přírodními jevy, na jevy společenské.

⁸ Toto dílo, jehož název lze do češtiny přeložit „Řád smyslovosti“, bylo publikováno v roce 1952.

aktivit) a na zásadní úlohu jeho vymáhání, pokud jde o prvky, které se daným pravidlům nepodřídí.⁹ Možnost svobodně využívat své individuální znalosti pro své vlastní účely a odpovědnost (vůči pravidlům),¹⁰ jakožto neoddělitelná součást této svobody,¹¹ jsou dva fundamenty přispívající k rozvoji tržního systému.

Pro přehlednost tedy shrňme, že v souladu s přístupem rakouské školy budeme tržním systémem rozumět tržní řád a vymezíme jej jako komplexní systém, který je z definice v detailech neuchopitelný. Tato komplexita – složitost – je dána jednak povahou systému samého, jedná se tedy o vlastnost systému per se, a jednak kognitivními omezeními hodnotitele (Řezníček, Smutný, 2012b), tedy jednotlivce, který je součástí daného systému.

Jako na spontánně se utvářející systém (sít' uzlů a vazeb nebudovaných vnější mocí) lze pohlížet i na jeden z nejvýznamnějších fenoménů informatizace, kterým je internet. Tento globální informační systém (Sklénák, 2001) se stal za poslední roky mocným médiem bez nadsázky ovlivňujícím životy milionů jednotlivců ve společnosti. Podstatou funkce tohoto komunikačního nástroje je, jak uvádí Josef Musil (2007, str. 31) nikoliv technika, nýbrž „organizační řešení, založené na myšlence celosvětové volně vytvářené sítě bez centra.“ Dopad internetu na člověka a společnost je možné a zároveň žádoucí zkoumat z mnoha různých perspektiv. Objevuje se zde tak nové významné transdisciplinární téma řešené odborníky z oblasti informatiky, ekonomie, psychologie, sociologie a dalších vědních oborů, přičemž jako zásadní se v oblasti aplikované informatiky jeví, jak poukazuje Nicholas Carr (2011) právě otázka po vlivu internetu na individuální práci s informacemi. V následující kapitole reflektující internet jako nástroj, který přispívá k rozvoji tržního systému, je argumentováno na základě výše presentovaných teoretických fundamentů.

Přispívá internet k rozvoji tržního systému?

V rámci internetu jsou poskytovány různé služby. Především se však jedná o komunikační prostředek a médium umožňující (ve srovnání s jinými) bezprecedentní datovou saturaci¹² a šíření informací,¹³ respektive, pomineme-li pro tuto chvíli otázku interpretace informačního obsahu disponibilních dat a roli individuální znalosti v tomto procesu (tedy schopnost se daného obsahu zmocnit),¹⁴ šíření do podoby dat formalizovaných znalostí. Jak bylo ukázáno výše, pro rozvoj tržního systému je důležité efektivní individuální rozhodování členů společnosti využívajících své individuální znalosti k dosahování svých cílů, kteří se zároveň musí ve svém jednání podříditi daným pravidlům.¹⁵ K vyšší efektivitě rozhodování napomáhá cenový systém a individuální znalost cen. Zde se jeví internet jako účinný nástroj, který je v praxi čím dál více využíván. Elektronický obchod¹⁶ významně usnadňuje komparaci cen zboží mezi

⁹ Ján Pavlík (2004b) k tomu uvádí, že pokud „schází moc a síla transcendentna a abstraktních ideologických principů, které donucovaly kdysi jedince k tomu, aby respektovali pravidla tržního řádu a utvářeli tak moderní civilizaci, zbývá již pouze jediné: sekulární moc důsledného a přísného legálního sankcionování dodržování oněch pravidel.“

¹⁰ Jak píše Hayek (1998, str. 48 a 51): „Spontánní řády vyplývají z toho, že se jejich prvky podřizují jistým pravidlům chování.“

¹¹ Připomeňme Hayekovo resolutní odlišování pravého a tzv. falešného liberalismu, který se naopak vyznačuje záměrem pravidla nerespektovat a překonávat. Hayek (1960, str. 71) k tomu uvádí: „Svoboda neznamená jenom to, že jedinec má jak příležitost, tak i tíhu volby; značí to rovněž, že musí nést důsledky svých činů a dostane se mu za to ocenění nebo ponese vinu. Svoboda a odpovědnost jsou od sebe neoddělitelné.“

¹² Což můžeme chápat jako příležitost i jako hrozbu.

¹³ Okamžité „celosvětové“ zpřístupnění aktualizované informace považují někteří (Musil, 2007) za nejdůležitější možnost internetu v oblasti informování.

¹⁴ Ke které se v dalším textu ještě dostaneme.

¹⁵ Neboť, jak poznamenává Hayek (1998, str. 51), „výsledný řád bude determinován právě prostřednictvím znalostí, jimiž jednotlivci disponují, a které se týkají jak pravidel, tak konkrétních faktů.“

¹⁶ Podíl jednotlivců ve společnosti, kteří využívají internet k nakupování, se stále zvyšuje. Ze zveřejněných statistik (Eurostat, 2010) vyplývá, že zatímco v roce 2005 nakupovalo (v průběhu posledních 12 měsíců uskutečnilo alespoň jednu objednávku pro soukromé účely) v České republice přes internet 15,5 % uživatelů internetu (ve věku 16-74 let), v roce 2009 realizovalo nákup prostřednictvím internetu již téměř 37 % jeho uživatelů. Jejich počet

konkrétními prodejci zveřejňujícími cenu daného zboží na svém webu. Existují rovněž tzv. cenové srovnávače, tedy služby, které za nás tuto komparaci provádějí. Náklady na získání informací (respektive znalosti) o cenách jsou tak výrazně sníženy a internet zde působí jednoznačně pozitivně, jde-li o časový faktor rozhodování. Využívání internetu rovněž umožňuje snazší orientaci nejen, pokud jde o cenu, ale díky sdílení zkušeností mezi zákazníky ve formě komentářů hraje roli také při výběru dané komodity dle jejích kvalitativních charakteristik. Otázkou je pak analyzování (vyhodnocování) těchto komentářů. Tolik tedy ve stručnosti pokud jde o orientaci individua v rámci tržního systému a její zefektivnění díky využívání internetu. Druhou podmínkou rozvoje tržního systému je dodržování daných „pravidel chování“, kde se jeví internet rovněž jako potenciálně mocný nástroj. Uveřejňování důležitých údajů ve volně a jednoduše přístupných informačních zdrojích má potenci působit směrem k vyšší transparentnosti. Může vytvořit tlak na dodržování pravidel a zamezit projevům libovůle, tedy chápání svobody jako absenci odpovědnosti.

Na straně druhé je nezbytné uvažovat otázku validity a kredibility skrze internet prezentovaných informací. S tím souvisí role individuální lidské znalosti (její význam) v procesu interpretace informace (Řezníček, Smutný, 2012a). Pokud nemáme dostatečné penzum znalostí o prezentované skutečnosti, v našem případě tedy např. o výrobku, jež si zamýšlíme pořídit, tak se při spolehnutí na uvedené zákaznické zkušenosti vystavujeme nebezpečí, které představují zkreslené a mylné informace mající za následek naší realitě neodpovídající znalost. V tomto případě je účelné nespolehnout se pouze na jeden či několik málo zdrojů, ale porovnat jich pokud možno více, případně zjistit dodatečné informace z dalších zdrojů (např. mimo ony zákaznické/uživatelské recenze). Nově nabytá znalost nám umožní interpretovat informace jako falešné, a tyto pak nebudou mít na naše rozhodování negativní dopad. V uvedeném kontextu je zásadní problém vzdělávání. To by mělo v reakci na změněné podmínky v přístupu k informačním zdrojům klást důraz na rozvoj (kritického) myšlení a na rozdíl od pozorovatelných procesů tendujících, jak poukazuje Konrad Paul Liessmann (2009), k averzi k porozumění a k upozadění tzv. pro praxi neužitečné obecné vzdělanosti,¹⁷ by se nemělo nechat oklamat pocitem informační saturace.

Závěr

Príspevek se na základě prezentovaného myšlenkového rámce, kde byly naznačeny některé fundamenty ekonomické teorie rakouské školy, jako je metodologický dualismus apriorismus a individualismus, stejně jako chápání ekonomiky jako komplexního ex definitione v detailech neuchopitelného systému, ve kterém je jednotlivec „omezeně racionální“, pokusil nastínit možné řešení v jeho názvu vyřčené otázky. Internet jako dnes významné a ve srovnání s jinými v mnohém specifické médium jistě nikoli zanedbatelně změnil přístup jednotlivců k informacím. Navíc se zdá být nástrojem pro realizaci procesů směřujících k rozvoji tržního systému, ke kterému podle rakouských teoretiků přispívá za prvé lepší informovanost zefektivňující individuální rozhodování a za druhé individuální respekt vůči pravidlům. Pokud jde o bod první, v textu bylo poukázáno na rozhodování zefektivňující možnosti internetu, který umožňuje snazší komparabilitu cen i kvalitativních charakteristik jednotlivých komodit. Pokud jde o bod druhý, bylo ukázáno, že využívání internetu jako média pro zveřejňování důležitých dat, které je s to bezprecedentním způsobem šířit, může pozitivně působit směrem ke zvýšení transparentnosti a tlaku na dodržování daných pravidel. Další vývoj bude patrně záviset na tom, zda přijde adekvátní reakce na tyto nové fenomény z pozice vzdělávacího systému, jelikož na nás bezprecedentní datová dostupnost klade nové nároky. Naneštěstí právě technologický pokrok nezřídka podporuje negativní procesy ve vzdělávání směřující k jeho instrumentalizaci a komodifikaci.

přitom rovněž roste, zatímco v roce 2005 bylo v ČR 2 791 000 uživatelů internetu (jednotlivců, kteří použili internet alespoň jednou za poslední 3 měsíce) starších 16 let, v roce 2013 bylo v dané socio-demografické skupině 6 048 000 uživatelů internetu (ČSÚ, 2013).

¹⁷ Lze citovat Liessmanna (2009, str. 51), který upozorňuje, že „externí faktory – trh, zaměstnanost, kvalita lokality a technologický rozvoj – jsou nyní ty standardy, kterým má vzdělanec dostát. Z tohoto hlediska se jeví všeobecné vzdělání stejně postradatelné jako rozvoj osobnosti. V rychle se proměňujícím světě, v němž se kvalifikace, kompetence a obsahy vědění údajně stále mění, se nevzdělanost, tedy rezignace na závazné duchovní tradice a klasické vzdělání, stala ctností, která jednotlivci umožňuje rychle, flexibilně a bez zatížení vzdělanostním balastem reagovat na neustále se měnící požadavky trhu... To, co se realizuje ve vědění společnosti vědění, je sebevědomá nevzdělanost.“

V současné době probíhá návrh výběrového šetření, které bude realizováno formou dotazníku. Toto šetření by mělo (coby metoda sběru primárních dat) ověřit v disertační práci formulované modely vysvětlující nejen souvislosti naznačené v tomto příspěvku, ale i další fenomény, tematicky spadající do autorem řešené problematiky významu a hodnoty lidské znalosti v kontextu informatiky a moderních informačních technologií aplikovaných v rámci socio-ekonomických procesů.

Literatura

CARR, N., 2011. *The Shallows: How the internet is changing the way we think, read and remember*. Paperback edition, London: Atlantic Books. ISBN 978-1-84887-227-1.

ČSÚ, 2013. *Uživatelé internetu ve věku 16 let a více v ČR*. Šetření o využívání ICT v českých domácnostech a mezi jednotlivci (VŠIT). [online]. [cit. 2014-01-20]. Dostupný z: <http://www.czso.cz/csu/redakce.nsf/i/telekomunikacni_a_internetova_infrastruktura>.

EUROSTAT, 2010. *Individuals purchasing over the internet for private use - 2005-2009*. Údaje z šetření „Community Survey on ICT Usage in Households and by Individuals“ zveřejněné Českým statistickým úřadem. [online]. [cit. 2013-12-20]. Dostupný z: <http://www.czso.cz/csu/redakce.nsf/i/domacnosti_a_jednotlivci>.

HAYEK, F. A., 1945. The use of knowledge in society. *American Economic Review*. September. vol. 35, no. 4, s. 519-530, [cit. 2013-12-6]. Dostupný z: <<http://www.econlib.org/library/Essays/hykKnw1.html>>.

HAYEK, F. A., 1955. Degrees of Explanation. *The British Journal for the Philosophy of Science* 6 (23): 209–225.

HAYEK, F. A., 1960. *The Constitution of Liberty*. The University of Chicago, Gateway Editions, Ltd., South Bend, Indiana.

HAYEK, F. A., 1995. *Kontrarevoluce vědy: Studie o zneužívání rozumu*. Praha: Liberální institut. Přel. J. Pavlík, D. Slouková. ISBN 80-85787-87-3.

HAYEK, F. A., 1998. *Právo, zákonodárství a svoboda: nový výklad liberálních principů spravedlnosti a politické ekonomie*. 2. vyd. - Praha: Academia. 415 s. Přel. Tomáš Ježek, ISBN 80-200-0241-3.

HAYEK, F. A., 2008. *Cesta do otroctví* (The Road to Serfdom). 2. vyd. Brno: Barrister & Principal. 216 s. ISBN 978-80-87029-32-9.

HAYEK, F. A., 2011. *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*. The University of Chicago Press. (originally published 1952), ISBN 0-226-32094-4.

HEYLINGEN, F., JOSLYN, C., 2004. *Kybernetika a kybernetika druhého řádu*. New York, Encyklopedia of Physical Science & Technology, 2001. Přel. David Průdek a upr. Petr Parma. [cit. 2013-12-11] Dostupný z: <<http://www.systemic.cz/document/cybernetics.pdf>>.

LISSMANN, K. P., 2009. *Teorie nevzdělanosti: omyly společnosti vědění*. Přel. Jana Zoubková. Praha: Academia. 127 s. Název originálu: Theorie der Unbildung. ISBN 978-80-200-1677-5.

MISES, L., 1969. *Theory and History: An Interpretation of Social and Economic Evolution*. 2. vyd. New Rochelle, N. Y.: Arlington House.

MISES, L., 2007. *Human action: a treatise on economics*. Indianapolis: Liberty Fund, 4 v. ISBN 97808659768184.

MUSIL, J., 2007. *Komunikace v informační společnosti*. 1. vyd. Praha: Univerzita Jana Amose Komenského. 144 s. ISBN 978-80-86723-39-6.

PAVLÍK, J., 2002. *Filosofické základy metodologie ekonomických věd I*. Praha: Oeconomica. ISBN 80-245-0468-5.

PAVLÍK, J., 2004a. *F. A. Hayek a teorie spontánního řádu*. 1. vyd. Praha: Professional Publishing. 805 s. ISBN 80-86419-57-6.

PAVLÍK, J., 2004b. *Hayek pro XXI. století*. Liberální institut [online]. [cit. 2013-10-05]. Dostupný z: <<http://www.libinst.cz/stranka.php?id=38>>.

ŘEZNÍČEK, V., 2011. Hayekova kritická reflexe hranic rozumu. *E-LOGOS: ELECTRONIC JOURNAL FOR PHILOSOPHY* [online]. [cit. 2013-09-08]. 21 s. ISSN 1211-0442. Dostupný z: <<http://nb.vse.cz/kfil/elogos/student/reznicek11.pdf>>.

ŘEZNÍČEK, V., SMUTNÝ, Z., 2012a. Informační (ne)gramotnost v informační společnosti. Role znalosti v procesu interpretace významu informace. *Media4u Magazine* [online], roč. 9, č. 1, s. 107–110. ISSN 1214-9187. Dostupný z: <<http://www.media4u.cz/mm012012.pdf>>.

ŘEZNÍČEK, V., SMUTNÝ, Z., 2012b. Problematika složitosti přirozených a umělých systémů. Praha 14. 11. 2012. In: *Systémové přístupy 2012* [CD-ROM]. Praha: Oeconomica. s. 51–56. ISBN 978-80-245-1921-0.

SCHWARZ, J. za kol., 2000. *HAYEK SEMPER VIVUS*. Praha: Liberální institut. ISBN 80-86389-03-0.

SKLENÁK, V., 2001. *Data, informace, znalosti a Internet*. Praha: C. H. Beck. ISBN 80-7179-409-0.

JEL Classification: A12, D80, D83, O33, O34.

Summary

Does the Internet contribute to the Development of Market System?

This paper, as its name suggests, deals with the question whether the internet as a significant phenomenon of informatization (global information system) contributes to the development of market system. We examine this phenomenon from the point of view of the austrian school (as school of economic thought), which is based on the methodological apriorism, dualism and individualism. This school understands the economy as a complex (spontaneous) system, which unlike human purpose-made artifacts (organizations), is not the product of deliberate activities of individuals (of human design), but arises as a result of unplanned egoistic activities of millions of individuals (of human action), who obey the defined rules. In accordance with a second-order cybernetics this school looks at a specific individual as a factor, which does not stand outside the system, but is its active part, which suffers, as an element of complex system, by limited knowledge of the facts. This theoretical perspective gives us an unique view on this phenomenon, especially in the context individuals accessing information under the influence of internet changed conditions. It turns out, that the internet enables, not only because of easier comparability of prices, more effective decision-making within otherwise complex system, but it may also be an important tool in helping to enforce mentioned „rules of conduct“. On the other hand we discuss the market and individual interest group influenced nature of the contemporary information sharing environment which has the potential to lead to dissemination of distorted and erroneous information. This represents an issues in the context of education and the role of human knowledge, because human knowledge seems to be having a crucial influence on the society, based on the future development forecasts.

Keywords: Information, knowledge, internet, market system, complex system, spontaneous order, economy.

Přehodnocení přístupů komerčních subjektů k řízení marketingových aktivit v prostředí služeb na internetu

Zdeněk Smutný
zdenek.smutny@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: doc. Ing. Prokop Toman, CSc., (prokop.toman@vse.cz)

Abstrakt: Předložený příspěvek má několik důležitých momentů a obsahově představuje hlavní náplň připravované disertační práce. Nejprve je představena řešená problematika z netradiční pozice subjektu a jeho prostředí, na jehož základě jsou identifikovány důležité atributy, jak na straně subjektu tak prostředí, které jsou podstatné z hlediska marketingových aktivit. Dále jsou identifikovány problémy, kterým čelí firmy při svých marketingových aktivitách v prostředí služeb na internetu. Z toho vychází i úvahy směrem k přehodnocení přístupu komerčních subjektů k řízení těchto marketingových aktivit.

V souvislosti s tímto tématem je zajímavou otázkou současné využívání sociálních médií (aktuální fenomén doby) komerčními subjekty či pocítovaná důležitost těchto nástrojů řídicími pracovníky. Z toho důvodu jsou představeny dva průzkumy – jeden evropský a druhý vlastní průzkum provedený na více než 4500 komerčních subjektech v České republice. Na základě těchto průzkumů můžeme konstatovat pokračující růst pocítované důležitosti i praktického využívání sociálních médií komerčními subjekty pro komunikaci.

Vzhledem k identifikovaným problémům a s přihlédnutím k socio-ekonomickým dopadům internetu jakožto fenoménu je představeno schéma metodiky pro podporu marketingových aktivit v prostředí služeb na internetu. Představovaná metodika je postavená na třech pilířích, jež vychází z teorie komplexních sítí, modelování socio-technické interakce a genetických algoritmů. Tato metodika by měla být hlavním přínosem připravované disertační práce.

Klíčová slova: internet, marketing, řízení, metodika, subjekt, interakce

Úvod

Jedním ze zásadních bodů úspěchu marketingových aktivit komerčního subjektu je jeho komunikace. Tato potřeba se ještě zintenzivnila s rozvojem informačních a komunikačních technologií (ICT). Z nových fenoménů co do komunikace je v současnosti stále významnějším internet (jakožto médium) a především služby, které jsou přes něj nabízeny (internet-based services). Vytváří se zde nový svěbytný prostor, který není pouze alternativou k tomu fyzickému, nýbrž je s ním také propojen. Tento prostor zprostředkovávaný internetem se z hlediska uživatele mění v závislosti na službě, která je daným uživatelem využívána. Každá taková služba (např. Facebook, Youtube aj.) nabízí vlastní možnosti, pravidla či zákony, ve kterých se její uživatel nachází (ať už jde o člověka nebo softwarového robota – bota). V tomto pohledu je tedy každá taková služba vlastním ohraničeným systémem – možným světem. Prostor propojených internetových služeb (např. sociální média) lze jako celek na základě těchto charakteristik nazvat virtuálním prostředím (srov. s definicí virtuální organizace nebo týmu). (Smutný, 2013) Uvedený termín lépe akcentuje neustálé vytváření nových služeb (možných světů), kteréžto mohou, ale nemusí mít odraz ve světě fyzickém, na rozdíl od často uplatňovaného názvu „kyberprostor“ (Lévy, 2001), (Skyttner, 2005, str. 453), jenž lze parafrázovat jako řízení tohoto prostředí člověkem (tvůrcem), nebo pojmu „internet“, který je fakticky pouze síťovým propojením počítačů.

Člověk je nejenom v roli tvůrce, ale také subjektem tohoto nového prostředí. Důležitým kvalitativním posunem je, že není sám, kdo je subjektem tohoto prostředí, neboť kupříkladu uměle vytvořený bot je schopen ekvivalentní interakce včetně komunikace (přestože percipuje dané prostředí jinak než člověk a samozřejmě nemá vědomí). V tomto prostředí tak dochází k doposud největšímu sblížení lidského a umělého aktéra na úrovni kognitivních možností (člověk je „kognitivně ořezán“ a umělý aktér je uzpůsoben danému prostředí). Tento trend je dále podporován směrem k Web 3.0 technologiím sémantizace webu a lze oprávněně předpokládat, že bude docházet ke stále většímu otevírání virtuálního prostředí umělým aktérům, stejně jako jejich možností interakce.

Jako výchozí postoj k hlubšímu pochopení této problematiky (prostředí a subjektu) je vhodné zmínit přístup filosofie, jež poskytuje základ skrze tematizaci vzájemné konstituce subjektu a jeho prostředí, viz spor o subjekt v průběhu 20. století a problematika jeho (re)determinace prostředím/strukturou/jazykem (Descombes 2004). Samotný koloběh redeterminace na obou úrovních, tedy jak agenta, tak prostředí, nalezneme také v enaktivním přístupu kognitivní vědy. (Varela, Thomson, Rosch, 1992) V tomto ohledu můžeme pokládat virtuální prostředí jako praktický výraz normativní implikaci filosofie. Dostáváme se tak na interdisciplinární úroveň, jež umožňuje obohatit přístupy uplatňované v informatice a marketingu o filosofické aspekty a s nimi související implikace či důležité momenty. (Smutný, Řezníček, 2012a) Tuto mezioborovou synergii se pokusím představit v následujícím odstavci.

Ve virtuálním prostředí dochází k multimodální interakci subjektu, ve které je skryta fragmentarizace subjektu v postmoderním slova smyslu. Jedná se o důležitou implikaci směrem k marketingovým aktivitám v tomto prostředí vycházející z tématu subjektu ve filosofii (směrem do fyzického prostředí). Můžeme tedy odhlédnout od klasického přístupu marketingu zaměřující se na člověka v omezené formě jednotky (marketing ve fyzickém prostředí), ale naopak zohledňovat jeho roztržistost podmíněnou užívanými službami, tedy prostředím. Na rozdíl od fyzického je ve virtuálním prostředí tato roztržistost více akcentována. Fragmentarizaci můžeme pozorovat například v rámci kvalitativně odlišných profilů jednoho člověka v různých službách (Facebook, LinkedIn, LibimSeTi, World of Warcraft aj.). Z tohoto důvodu je vhodné při komunikaci s potenciálním zákazníkem k němu přistupovat odlišně v rámci jednotlivých služeb, které používá. Dalším důsledkem filosofických aspektů problematiky subjektu je jeho sebekonstituce ve virtuálním prostředí. Umělý aktér tak může být dán do nového postavení, včetně využití tohoto pohledu při marketingových aktivitách. Již neplatí, že jediným subjektem hodným zřetele je člověk a umělé je chápáno pouze jako nástroj. Tento pohled je bohužel až na výjimky (optimalizace stránek pro vyhledávače ~ např. googleboty) preferován.

Vedle pochopení subjektu a jeho vztahu k prostředí je stejně důležitý také pohled a možnosti, které v souvislosti s marketingem toto prostředí přináší. S prudkým nárůstem počtu uživatelů připojených k internetu a jeho dostupností i v domácnostech začíná virtuální prostředí plně konkurovat tomu fyzickému. Zásadním problémem je jeho rychlý rozvoj a dopad na socio-ekonomické aktivity ve společnosti nutící marketingové specialisty neustále měnit strategii komunikačních kampaní v závislosti na současných trendech v internetových službách (v globálním i lokálním pohledu). Přichází tedy potřeba neustálého vzdělávání těchto pracovníků, a to nikoli pouze na úrovni schopnosti ovládnutí jednotlivých služeb jakožto nástrojů komunikace (např. Twitter), nýbrž na úrovni znalosti atributů virtuálního prostředí, jež jsou dány jeho konstitucí na softwarové i hardwarové úrovni, a na základě toho porozumět podstatě potenciálního nástroje pro marketingovou komunikaci a jeho adekvátního využití při strategickém plánování marketingových kampaní. (Smutný, Řezníček, 2012b; Smutný, Řezníček, 2013)

Tyto výchozí úvahy budou ještě krátce shrnuty v souvislosti s řešenou disertační prací v části: Přehodnocení pohledu na řízení marketingových aktivit v prostředí služeb na internetu. Cílem disertační práce je představit metodiku pro podporu řízení marketingových aktivit ve virtuálním prostředí. Cílem metodiky bude zefektivnit vynakládané finanční prostředky na marketingové aktivity v tomto prostředí s ohledem na dosažené výsledky napříč službami. Připravovaná metodika v sobě bude zohledňovat výše řečené a stojí na následujících pilířích:

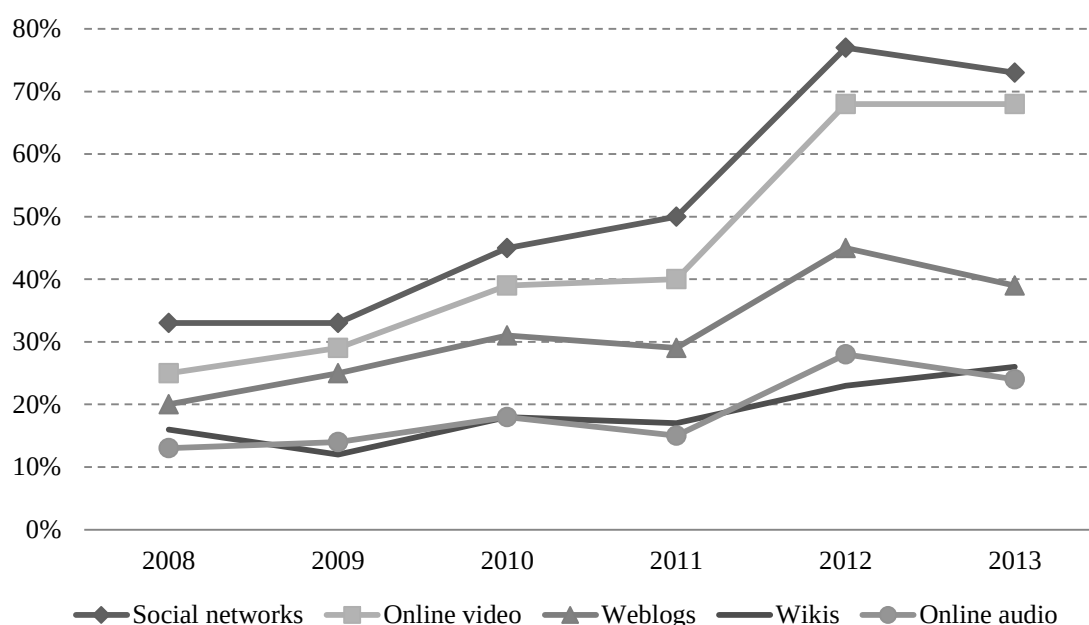
- *Využit vlastností internetu jakožto propojených počítačových sítí* a důsledky na jím zprostředkované prostředí služeb. Důraz bude kladen na možnosti využít tyto vlastnosti komplexní sítě (vycházející z teorie grafů) při řízení marketingových aktivit (např. problematika Influencer Marketing). Zároveň se jedná o výchozí pensum znalostí využitelných většinou napříč službami.
- *Zachytit a využít interakce* v tomto prostředí na úrovni subjektů (člověk, bot) a jejich dopad na marketingové aktivity v daném prostředí. Důraz bude kladen na úpravu dostupných metodik k modelování interakcí z hlediska sociální informatiky (na úrovni socio-technické interakce; člověk a ICT; tedy např. STIN – Socio-Technical Interaction Networks), jejich rozšíření do internetem zprostředkovaného prostředí a možnosti jejich využití v marketing managementu.
- *Využití přístupů umělé inteligence pro kontinuální vyhodnocování rentability jednotlivých marketingových aktivit v tomto prostředí*. Důraz bude kladen zejména na možnosti využití evolučních výpočetních technik pro optimalizaci především komunikačních kanálů.

Než budou podrobněji probrány základní principy, tak se pokusíme odpovědět na otázku: Jaký je současný přístup komerčních subjektů k sociálním médiím? Na tuto otázku se podíváme z různých úhlů pohledu (jak na úrovni vnitřní komunikace, tak jejich využití při propagaci) včetně s tím související diskuze. Takový vhled je v kontextu práce žádoucí, neboť odráží potřebnost řešení problematiky řízení marketingových aktivit v tomto prostředí, kterou si uvědomují i komerční subjekty. V navazující části bude načrtnuto základní schéma připravované metodiky. Na této metodice v současné době jako doktorand pracuji v rámci své disertační práce. Závěrem uvedu krátké pozastavení nad navazujícím problémem provázanosti fyzického a virtuálního prostředí směrem k implikacím do marketingu.

Vývoj přístupu evropských firem k sociálním médiím

V předchozí části byla představena východiska úvah o nutnosti změny pohledu na internet jako důležitý fenomén současné doby a motivace ke změnám na úrovni řízení marketingu. Směrem k řízení marketingových aktivit bylo odkázáno na nové přístupy k řešení problematiky řízení marketingové komunikace v prostředí služeb na internetu. S tím souvisí také posun u komerčních subjektů, které musejí zohledňovat celospolečenské změny (trendy) v souvislosti se socio-technickou interakcí (přijetí či zavrnutí určitých technologií na úrovni aktivního využívání). V tomto kontextu by bylo rovněž zajímavé podívat se na současný přístup komerčních subjektů k problematice komunikace a využívání sociálních médií, které tvoří páteř marketingové činnosti v tomto prostředí.

Dle evropského průzkumu (Zerfass et al., 2013), který byl publikován v červnu 2013, můžeme nalézt tento odraz ve vývoji přístupu řídicích pracovníků (starajících se o komunikační aktivity komerčních subjektů) k sociálním médiím a jejich důležitosti z hlediska komunikačních aktivit (Obrázek 1). Tento růst důležitosti sociálních médií v uplynulých šesti letech odráží úspěšnou penetraci nových technologií (na bázi ICT) ve společnosti a nárůst jejich užívání.

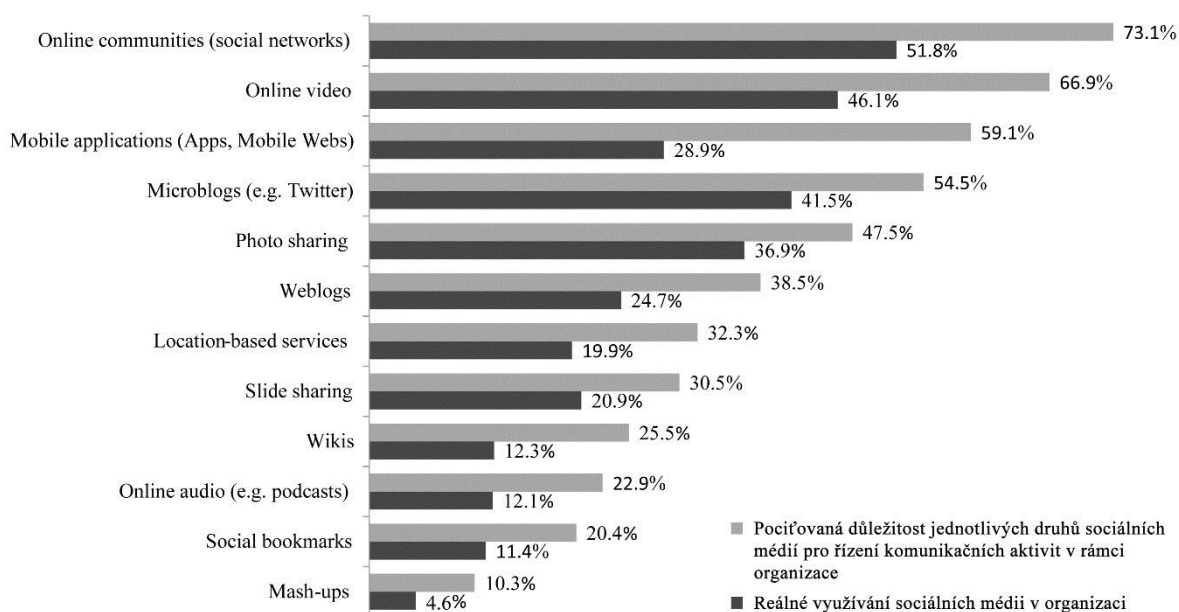


Obrázek 13: Vývoj vnímané důležitosti jednotlivých druhů sociálních médií v průběhu let 2008-2013.

Průzkum byl proveden na vzorcích 1542 až 2358 respondentů podle roku uskutečnění

Zdroj: Zerfass et al., 2013

Druhý obrázek ukazuje pocítovanou důležitost a reálné využívání sociálních médií jako alternativy za účelem zabezpečení komunikace uvnitř jednotlivých firem. Obrázek dobře ilustruje aktuální trendy v komunikaci. Především rozšíření využívání sociálních sítí, online videa a mobilních aplikací pro firemní komunikaci. Nicméně v tomto případě se jedná o pohled především na vnitřní komunikaci v rámci organizace, případně mezi dalšími organizačními celky.



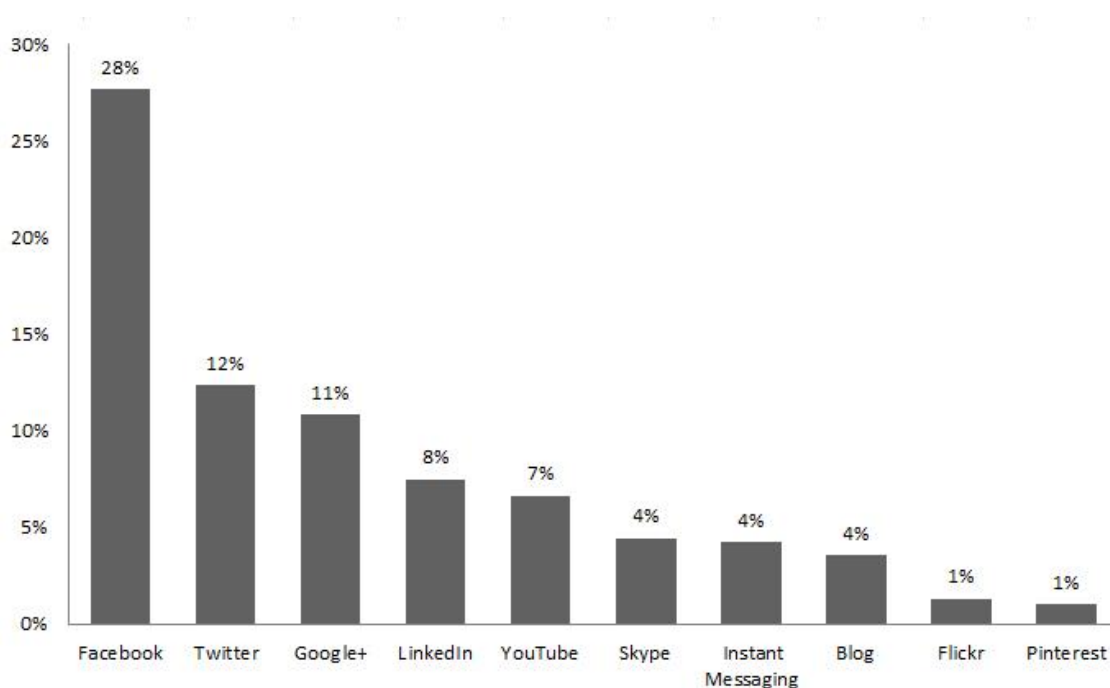
Obrázek 14: Firemní komunikace skrze alternativní nástroje a kanály v Evropě – pocitovaná důležitost versus reálné využívání v dané firmě. Vzorek o velikosti 2358 respondentů Zdroj: Zerfass et al., 2013

Kromě komunikace je z hlediska marketingového mixu zajímavá také oblast Distribuce. Produkty či služby jsou stále častěji nabízeny přes virtuální tržiště globálně (AOL Video, Apple AppStore, Google Play, Steam aj.). Jedná se o nejpřístupnější možnost globální distribuce digitálního zboží. Tento posun včetně historických konsekvencí je popsán v (Ya-Ping, 2012; Bhattacharjee et al., 2011). Uživatelé si na tento přístup již navykli a využívají jej stále více – například Apple oznámil, že za rok 2013 utratili jeho zákazníci přes AppStore¹ v přepočtu více než 200 miliard Kč, přičemž jen v prosinci 2013 bylo staženo přes 3 miliardy aplikací. Od spuštění tohoto tržiště v létě 2008 do dneška dochází neustále k nárůstu těchto čísel. (Apple, 2014) Tento trend potvrzuje i prognóza (Revenue Architects, 2012) v oblasti prodeje digitálního zboží (např. software), kde je odhadováno, že do roku 2015 bude 70 % prodeje a plateb uskutečněno online. Stvrzuje se tak pozitivní přijetí nových technologií společností, a to ať už na softwarové či hardwarové bázi. V důsledku toho jsou společnosti zaměřující se na digitální zboží nuceny měnit pohled také na celý marketingový mix a kontinuálně se přizpůsobovat vývoji a změnám (v tomto případě změnám na úrovni distribučních kanálů).

Využívání sociálních médií českými firmami

V tomto kontextu by bylo vhodné věnovat se také situaci v České republice. Bohužel aktuální, ucelené a volně dostupné průzkumy v rámci komerčních subjektů podnikajících v prostředí služeb na internetu schází. Z tohoto důvodu byl proveden vlastní průzkum s cíle zmapovat aktuální stav využívání sociálních médií komerčními subjekty, především na úrovni marketingové komunikace. Tento průzkum bude podrobněji vyhodnocen v připravovaných konferenčních příspěvcích v roce 2014. Na následujícím obrázku je vidět využívání vybraných sociálních médií procentuálně z celkového počtu 4584 sledovaných společností. Je třeba upozornit, že bylo sledováno, zda existuje propojení výchozí webové stránky společnosti s profily na jednotlivých službách – jediné takto lze zahrnout sociální média do integrované marketingové komunikace.

¹ V současnosti nabízí tento obchod přes milion aplikací pro iPhone, iPod a iPad, které jsou přes toto virtuální tržiště distribuovány do 155 zemí světa. (Apple, 2014)



Obrázek 15: Využívání vybraných sociálních médií českými firmami podnikajícími v prostředí služeb na internetu (Eshopy, Mediální agentury, On-line služby, Softwarové firmy, Vývojáři webu). Vzorek zahrnuje 4584 firem. Výzkum byl proveden na podzim 2013. Zdroj: Autor

Z průzkumu vyplývá, že konkurence mezi českými firmami z hlediska využívání sociálních médií pro aktivní propagaci je nízká. Výsledky můžeme srovnat s dříve provedeným průzkumem (Smutny et al., 2013a) v rámci Top 100 firem v USA a ČR (bez ohledu na druh podnikání). Zde bylo dosaženo podobných výsledků využívání Facebooku (29 %), nicméně v USA jej využívá 74 % firem. U dalších sledovaných sociálních médií (Twitter, YouTube, LinkedIn, Google+, Flickr, Pinterest) byl tento rozdíl i více než desetinásobný ve prospěch firem z USA. Závěrem lze konstatovat, že sociální média jsou z hlediska marketingového pohledu na internet důležitým fenoménem současnosti z důvodu jejich socio-ekonomických dopadů na společnost (např. každodenní využívání lidmi).

Přehodnocení pohledu na řízení marketingových aktivit v prostředí služeb na internetu

Nutnost řešit problematiku marketingu v prostředí služeb na internetu nastala společně s dostupností připojení k internetu (exponenciální růst se strmým nástupem kolem roku 2000), zvyšováním počítačové gramotnosti, finanční dostupnosti ICT a jeho penetrace do socio-ekonomických aktivit společnosti. Z hlediska firem byla důležitá oblast podnikání a velikost dané společnosti, kdy větší a nadnárodní společnosti začaly nový komunikační kanál využívat dříve. V současnosti již problém efektivního marketingu v tomto prostředí musí řešit většina firem, ať už se jedná o jakkoli velkou společnost napříč různými oblastmi zájmu. Možnosti marketingových aktivit v tomto prostředí z hlediska 4P se v uplynulých patnácti letech značně proměnily. V posledních letech se tato revoluce konala na úrovni nových služeb (např. Facebook, Twitter) a tvorby nových ekosystémů (např. v rámci společností Apple a Google), které byly kladně přijaty ve společnosti.

Zatímco ze začátku bylo pro marketingové pracovníky podstatné ovládnout novou službu z hlediska jejího uchopení nejčastěji za účelem komunikace, dnes je třeba posunout důraz na holistickou úroveň řízení marketingových aktivit v globálním virtuálním prostředí napříč užívanými službami. Důvodem je především stoupající počet služeb, které dnes běžně člověk využívá. Dnes se již nejedná o výzvu především pro větší společnosti, ale je to realita i u začínajících firem (minimálně na úrovni zajištění odrazu komerčního subjektu ve virtuálním prostředí – např. existence květinářství v internetových katalogích). Tyto trendy jsou zřetelné i ve výše uvedených výzkumech, kdy komerční subjekty chtějí být v daných službách vidět a komunikovat se svými (potenciálními) zákazníky.

Schéma navrhované metodiky pro podporu řízení marketingu

Vzhledem k těmto proměnám je vhodné začít u člověka, který připravuje kampaně a využívá za tímto účelem dané kanály/služby. Důraz je tedy třeba klást na znalosti, jež jsou obsaženy v podstatě prostředí budovaného na bázi ICT. Tento pohled je odlišný od upřednostňovaného zaměření pouze na (jednotlivé) nástroje a schopnost pracovníka je „ovládnout“ a používat. Cílem je, aby si marketingový pracovník byl vědom různých druhů interakcí a synergických efektů, ke kterým zde dochází. Proto je třeba jej za účelem řízení komunikačních aktivit v tomto prostředí vybavit výchozími znalostmi, na základě kterých provádí vyhodnocení dat získaných z prostředí. V rámci připravované metodiky se jedná o vlastnosti komplexních sítí spolu s vyhledáváním důležitých bodů/hubů s větším interakčním potenciálem, a to jak napříč službami, tak v rámci jednotlivých služeb (na bázi metodiky STIN).

Na základě těchto znalostí provádí pracovník výchozí volbu portfolia služeb (mix sociálních médií), jejich propojení s výchozí webovou stránkou (např. web společnosti) a nabídnutí dalších forem sdílení. Tento postup umožňuje sledování interakcí na třech úrovních – web (např. pomocí Google Analytics), jednotlivé služby v rámci portfolia (např. pomocí implicitních statistik) a sdílení napříč službami (např. pomocí AddThis). Na základě tohoto přístupu má marketingový specialista dostatek dat vhodných k interpretaci dosažených výsledků v širším kontextu včetně interakcí napříč službami, které jsou hůře postižitelné. Celý tento postup je vystavěn na zdarma dostupných technologiích. Tato problematika již byla představena v příspěvcích: (Smutný, Řezníček, 2012b; Smutný et al., 2013a) a směrem ke specifickému využití na úrovni řízení propagace (Smutný, 2014).

Dalším problémem, který je třeba řešit, je kontinuálního vyhodnocení získaných dat, efektivní přerozdělování finančních prostředků mezi jednotlivé služby (reklama, PR články, soutěže aj.) a případná implementace tohoto přístupu do existujících frameworků (rámců) uplatňovaných při řízení marketingu (např. Marketing communications planning Framework, Pragmatic Marketing Framework, RACE: Digital Marketing Improvement Framework). Důvodem pro nutnost kontinuálního vyhodnocování je turbulentní prostředí, ve kterém dochází k častým změnám – ať už na úrovni interakcí nebo preferencí služeb. Konečným cílem je vytvoření metodiky pro podporu řízení marketingu, která bude použitelná sama o sobě stejně jako součást existujících přístupů.

Problém kontinuálního vyhodnocování jsem se rozhodl vyřešit iteračním přístupem a optimalizací využívání jednotlivých kanálů za pomoci evolučních výpočetních technik – genetického algoritmu. Cílem je vybrat takový mix kanálů (portfolio služeb), který bude oslovovat co největší počet recipientů za co nejnižší cenu při maximalizaci pozitivních interakcí, ke kterým dochází v tomto prostředí (sdílení či přebírání zprávy (lidským i umělým aktérem), zhlédnutí zprávy, nákup, přechod na mateřský web aj.). Podívejme se nyní na Obrázek 4, na kterém bude představen základní princip. Níže pak bude uveden tento postup na jednoduchém příkladu.

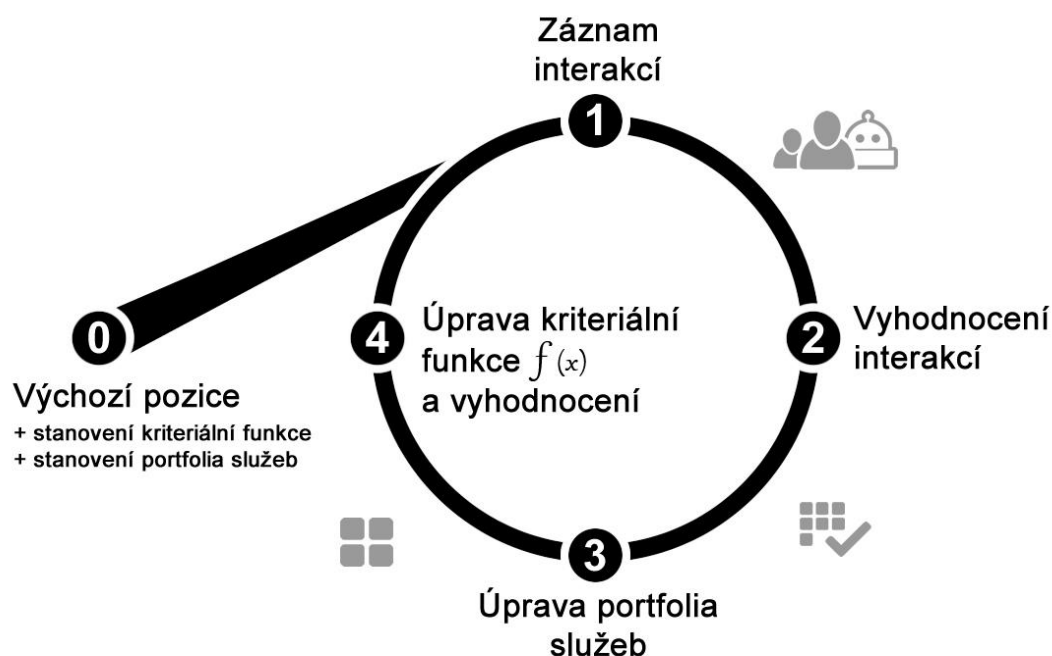
V **bodě 0** se provede expertní výběr portfolia služeb marketingovým specialistou (expert). Je především na něm vybrat co nejrelevantnější množinu služeb (např. portály, blogy, diskuzní fóra, speciální účty na sociálních médiích s velkým počtem příznivců), ze které se bude následně vybírat výchozí mix. Celková cena za využití těchto služeb by měla být vyšší než objem přidělených finančních prostředků určených pro jejich zaplacení. U každé služby je na základě komunikace s provozovatelem zjištěn počet potenciálně oslovených a cena. Pokud existuje více možností jak se dohodnout, tak by měl expert rozhodnout, jaký způsob spolupráce vzhledem k cílům aktuální kampaně bude pro danou službu nejlepší.

Řekněme, že expert vybral 32 služeb a dále spočítal celkový počet oslovených a celkový objem finančních prostředků, které by musely být vydány, kdyby mohl zaplatit za všechny služby. Každá služba má tedy procentuálně vyjádřený podíl na celkovém dosahu (oslovení) a procentuální podíl na finanční náročnosti (např. 0,11 oslovených za cenu 0,09). Zároveň je stanovena hranice maximální ceny – maximální objem finančních prostředků opět ve formě procentuálního podílu z celkové ceny za všechny služby, kupříkladu 0,56.

Dále je třeba vytvořit jedince v naší populaci, což může být náhodný výběr šesti služeb, přičemž délka chromozomu bude dvanáct (vybrané služby mají na i -tém genu 1 v opačném případě 0). Opakováním tohoto postupu budou vygenerováni další jedinci a tedy celá nultá populace. Pro vytvoření další generace je potřeba zvolit kritériální funkci, která bude vybírat nejvhodnější jedince na základě jejich zdatnosti.

Jejím cílem bude najít co nejvyšší počet oslovených za cenu zespoda nejbližší té maximální, tedy 0,56. Na základě toho budou vybráni nejvhodnější jedinci a pomocí jejich mutace bude populace doplněna. Tímto postupem se dojde k výchozímu mixu portfolia sedmi služeb, což bude kupříkladu jedinec o zdatnosti 0,86 oslovených za cenu 0,55.

Abychom zjistili, jak se tento výchozí mix osvědčí směrem k marketingovým cílům, je potřeba získat zpětnou vazbu ~ reakce na probíhající kampaň. To získáme v **bodě 1** záznamem proběhlých pozitivních interakcí za určité období (ohrazenou kampaň). Interakcí může být více typů a my musíme každou interakci ohodnotit vzhledem k marketingovým cílům. Nejtěžší však je přiřadit, jaká příčina způsobila určité interakce – **bod 2**. To záleží a zároveň je limitováno možnostmi získání dat o interakcích daným expertem. Řekněme, že expert bude upřednostňovat všechny formy sdílení a diskuzí, které bude hodnotově upřednostňovat. Každá služba tak dostává třetí parametr pro hodnocení, a tím je pozitivní interakce opět ve formě procentuálního podílu na celkové interakci. Při vyhodnocování může expert narazit na tzv. huby, kde dochází k většímu množství interakcí (může se jednat jak o využívané služby, tak doposud nevyužité). V **bodě 3** expert může vyřadit očividně neúspěšné služby bez zpětné pozitivní reakce, nebo ty, kde došlo k negativní reakci. Zbylé služby doplní nově identifikovanými huby nebo novými službami – např. kvůli úpravám v marketingových cílech (jiná cílová skupina).



Obrázek 16: Náčrt metodiky pro podporu marketingového řízení aktivit komerčních subjektů v prostředí služeb na internetu. Zdroj: Autor

Tímto způsobem byl určen nový mix služeb, kdy můžeme jeden parametr u každé služby upravit tak, že využijeme data z proběhlé kampaně a využijeme reálnější počty oslovených naší akcí. Kritériální funkce v **bodě 4** bude nyní preferovat jedince, kteří budou mít co nejvyšší hodnoty oslovených a pozitivních interakcí a jejich cena zespoda nejbližší té maximální (na další iteraci může být přidělen jiný objem finančních prostředků). Kritériální funkci můžeme upravit směrem k novému objemu budgetu.

Kontinuální vyhodnocování je důležité i z toho důvodu, že se mění zaměření kampaně dle aktuálních preferencí firmy (např. image vs. prodej, úprava cílové skupiny), kdy dochází k jiným interakcím, a proto je třeba se zaměřit na služby, které tyto interakce přinášejí nejvíce. Obdobně se mění i množství finančních prostředků (před Vánocemi bývá budget vyšší). Výše uvedená metodika není definitivní a bude ještě upravována, tak aby mohla být co nejjednodušeji aplikována v praxi. Zásadní nevýhodou je silný lidský činitel, na kterém spočívá i finální rozhodnutí, zda nabízená možnost je opravdu ta správná a nedošlo k opomenutí nějakého důležitého faktoru. Na druhou stranu, v případě rozsáhlých kampaní napříč státy světa může tato metodika být přínosná ve smyslu utřídění priorit a rychlého nabídnutí optimálního mixu.

Propojování komunikačních nástrojů ve fyzickém a virtuálním prostředí

Zajímavou související problematikou je souběh fyzického a virtuálního prostředí. I když je výše navržená metodika zaměřená čistě pro podporu řízení marketingových aktivit v prostředí služeb na internetu, zajímavým rozšířením pro některé firmy by bylo její využití na úrovni, kde se propojuje fyzické a virtuálních prostředí. V následující části stručně představím příklad tohoto propojování na úrovni marketingové komunikace.

Při vytváření strategie komunikační kampaně a tvorbě mediaplánu je cílem marketingového odborníka maximalizovat komunikační potenciál jednotlivých médií jejich paralelním zapojením tak, aby byla kampaň intenzivní a aby docházelo u jejích recipientů k opakovanému diferenciovanému sdělení přes různá média. Kombinací jednotlivých komunikačních nástrojů vzniká synergie, kdy recipient vnímá celek jako jeden tok informací z vícera zdrojů. Jedná se o integrovanou marketingovou komunikaci, která může pokrývat jak fyzické, tak virtuální prostředí.

Do tohoto vidění je tedy třeba promítnout nejenom zpravidla jednosměrné komunikační nástroje užívané ve fyzickém prostředí, ale také zapojit interaktivní nástroje virtuálního prostředí. Zřejmě nejzajímavější způsob propojení kampaní ve fyzickém a virtuálním prostředí představují dnes možnosti mobilních zařízení. Nejčastěji se tak děje formou odkazování buď formou URL odkazu nebo přímého ale stále méně známého QR kódu. Jejich cílem je co nejrychleji a nenáročně přesunout recipienta sdělení ve fyzickém světě do virtuálního prostředí – kupříkladu na produktový/firemní profil na sociální síti, web, případně vyvolání specifické akce jako stažení aplikace z virtuálního tržiště (Google Play, Apple AppStore aj.).

Obdobně funguje i propojení z druhé strany, kdy určitá softwarová aplikace může nabídnout svému uživateli odkaz ve formě zeměpisných souřadnic, na jejichž pozici dojde k další akci – příkladem může být hra pro věrné zákazníky na bázi geocachingu.

Na jedné straně musí být marketingový specialista seznámen s aktuálně dostupnými technologiemi (v mobilních zařízeních), na druhé straně musí být recipient sdělení schopen a chtít danou technologii používat (takto interagovat). Na základě svých znalostí (myšlení a kreativity) může marketingový pracovník vhodně kombinovat komunikační nástroje za účelem maximalizace jejich synergie. Výsledkem je pak ucelená integrovaná marketingová komunikace – jak ve fyzickém tak virtuálním prostředí.

Závěr

Předložený příspěvek shrnuje můj dosavadní pokrok směrem k disertační práci, tedy tvorbě metodiky pro podporu řízení marketingu v prostředí služeb na internetu. Je zřejmé, že se jedná v současnosti o prakticky orientovanou výzvu, které musí firmy čelit, a to především z důvodu své konkurenceschopnosti v tomto prostředí. Nadále bude pro moji práci nutné rozvíjet a konsolidovat myšlenku kontinuálního vyhodnocování a optimalizace využívaných (komunikačních) kanálů a rozprostření finančních zdrojů mezi ně. Výslednou metodiku bude také nutné vhodně zařadit do existujících širších rámců řízení používaných v oblasti marketingu a na závěr ji pilotně otestovat.

Přidanou hodnotou tohoto příspěvku je především vlastní průzkum na vzorku 4584 firem v České republice, za účelem zjištění: Jaká využívají sociální média pro komunikace se zákazníkem. Jiný pohled v kontextu využití sociálních médií pro vnitrofiremní komunikaci, je představen v souvislosti se studií (Zerfass et al., 2013).

Na závěr bych také chtěl poukázat na dosavadní publikační výstupy především na mezinárodních konferencích, kde byla tato problematika prezentována a diskutována. Na základě těchto výsledků jsem s kolegy podal v roce 2013 žádost o interní grant s názvem: Inovativní pohled na hodnotu zákazníka a další faktory ovlivňující řízení marketingu. Začátkem roku 2015 bych rád publikoval kompletní metodu pro podporu marketingového řízení včetně prvních výsledků z případové studie.

Literatura

- APPLE, 2014. Prodeje v obchodě App Store v roce 2013 dosáhly 10 miliard USD. In: *Apple, Inc.* [online]. [cit. 2014-01-21]. URL: <http://images.apple.com/cz/pr/library/2014/01/07App-Store-Sales-Top-10-Billion-in-2013.pdf>
- BHATTACHARJEE, Sudip et al., 2011. Digital Goods and Markets: Emerging Issues and Challenges. *ACM Transactions on Management Information Systems*, roč. 2, č. 2. DOI: 10.1145/1985347.1985349
- DESCOMBES, Vincent, 2004. *Le complément de sujet: enquête sur le fait d'agir de soi-même*. Gallimard, 2004, 521 s. ISBN 20-707-6130-4.
- LÉVY, Pierre, 2001. *Kyberkultura*. Karolinum, Praha. ISBN 80-246-0109-5
- REVENUE ARCHITECTS, 2012. Software Distribution channel: The Reports of its death are greatly exaggerated. [cit. 2014-01-25]. Available from: <http://www.avangate.com/lp/whitepaper-software-distribution-channel-management-2012.html>
- SKYTTNER, Lars, 2005. *General systems theory problems, perspectives, practice*. 2. vyd. Hackensack, NJ. ISBN 98-127-7475-0.
- SMUTNÝ, Zdeněk, ŘEZNÍČEK, Václav, 2012a. Příspěvek k současné filosofii a metodologii vědy se zaměřením na inter- a transdisciplinární výzkum. *E-Logos – Electronic Journal for Philosophy* [online], roč. 3, č. 3, s. 1–21. ISSN 1211-0442. URL: <http://nb.vse.cz/kfil/elogos/science/smutny12.pdf>.
- SMUTNÝ, Zdeněk, ŘEZNÍČEK, Václav, 2012b. Zvyšování efektivity marketingové komunikace ve virtuálním prostředí. In: *Nové trendy v marketingu – Zodpovednosť v podnikaní*. Trnava: FMK UCM, 2012, s. 196–204. ISBN 978-80-8105-439-6.
- SMUTNÝ, Zdeněk, 2013. Socio-technická interakce: virtuální prostředí jako důsledek informatizace společnosti. In: *Kognitivní věda a umělý život*. Opava: Slezská univerzita, s. 259–294. ISBN 978-80-7248-863-6.
- SMUTNÝ, Zdeněk, ŘEZNÍČEK, Václav, 2013. Role of human knowledge in contemporary information society and its individual nature. In: *Active Citizenship by Knowledge Management & Innovation*. Zadar: ToKnowPress, s. 453–459. ISBN 978-961-6914-01-7. ISSN 2232-3309.
- SMUTNÝ, Zdeněk et al., 2013a. Interaction of Social Media and Its Use in Marketing Management. In: *IDIMT-2013 Information Technology Human Values, Innovation and Economy*. Linz: Trauner Verlag, s. 167–174. ISBN 978-3-99033-083-8.
- SMUTNÝ, Zdeněk et al., 2013b. Případová studie užití Twitteru pro propagaci: Korelace mezi množstvím zveřejněných Tweetů a zpětnou vazbou od uživatelů Twitteru. *ProInflow* [online], č. 2, s. 1–9. ISSN 1804-2406. URL: <http://pro.inflow.cz/sites/default/files/pdfclanky/twitter.pdf>.
- SMUTNÝ, Zdeněk, 2014. Environment-driven marketing communication. In: *12th International Conference Hradec Economic Days*. Hradec Králové: Gaudeamus.
- VARELA, Francisco J., Evan THOMPSON a Eleanor ROSCH, 1992. The embodied mind: cognitive science and human experience. 7. vyd. Cambridge: MIT Press. ISBN 978-026-2720-212.
- YA-PING, Hu, 2012. E-Marketing Development in Virtual Market-Space: A Strategic Perspective. *Asian Journal of Business Management*, roč. 4, č. 4, s. 359-366. ISSN: 2041-8752
- ZERFASS, Ansgar et al., 2013. *European Communication Monitor 2013. A Changing Landscape – Managing Crises, Digital Communication and CEO Positioning in Europe. Results of a Survey in 43 Countries*. Brussels: EACD/EUPRERA, Helios Media. ISBN 978-3-942263-21-4.
- JEL Classification:** M13, M31, M37.
- Dedikace:** Příspěvek je zpracován jako součást výzkumného projektu IGA, VŠE v Praze: Inovativní pohled na hodnotu zákazníka a další faktory ovlivňující řízení marketingu.

Summary

Rethinking approaches of commercial subjects to managing of marketing activities in the environment of services on the internet

This paper has several important moments and its content is the main fulfilment of the prepared dissertation. First is presented issue from an unusual position of a subject and its environment, on this basis are identified important attributes for both the subject and the environment that are significant from the viewpoint of marketing activities. There are identified problems faced by companies in their marketing activities in the environment of services on the internet. This is based for considerations direct to rethinking approaches of commercial subjects to managing of marketing activities.

In the context of this topic are an interesting issue the current using of social media (as actual phenomenon) by commercial subjects and the perceived importance of these tools by managers. For this reason are introduced two surveys – one European and second is own survey on more than 4500 companies in the Czech Republic. Based on these surveys, we claim the continuing growth of the perceived importance and practical use of social media by commercial subjects for communication.

Due to the identified problems and taking account of socio-economic impacts of the internet as a phenomenon is introduced scheme of methodics to support marketing activities in the environment of services on the internet. The presented methodics is based on three pillars, which are based on the theory of complex networks, modelling of socio-technical interaction and genetic algorithms. This methodics should be the main benefit of the prepared dissertation.

Keywords: internet, marketing, management, methodics, subject, interaction.

Metriky, monitoring a řídicí proces Data Governance

Martin Vacek
xvacm31@vse.cz

Doktorand oboru Informatika

Školitel: doc. Ing. Jan Pour, CSc.

Abstrakt: Článek se zabývá identifikací a definicí důležitých oblastí Data Governance, kterými jsou metriky a procesy se zaměřením na procesy řídicí. Na základě analýzy české a světové literatury a praktických zkušeností autora je nejprve definován informační rámec, s kterým autor pracuje, včetně definicí klíčových pojmů (Data Governance, Datová kvalita, Data Governance Council aj.) a dále je popsán aktuální stav Data Governance ve světě. Zmíněny jsou konkrétní výzvy, kterým odborníci na řízení dat v současnosti čelí. Patří mezi ně nízká víra zainteresovaných osob v přínosy Data Governance a nové požadavky regulačních úřadů. Tyto problémy se týkají organizací bez ohledu na daný sektor. Postupně jsou pak popsány metriky užívané pro řízení dat v rozdělení po jednotlivých vybraných objektech řízení (data, role, procesy). Následuje kapitola zaměřená na popis monitoringu a reportingu zmíněných metrik. Monitoring je zde charakterizován především z pohledu času a z hlediska automatizace. Popsán je aktuální trend reportingu podávat informace o datech odpovědným osobám efektivně prostřednictvím tzv. Data Governance Dashboards, s čímž souvisí i další část kapitoly, která se věnuje nastavování úrovně detailu měřených metrik a jejich agregace pro různé úrovně řízení resp. útvary Data Governance. Nakonec je definován klíčový řídicí proces Data Governance. Pozornost je zde věnována i řízení financí a v neposlední řadě také rozdílům v přístupech k řízení dat v různých sektorech ekonomiky a u různých velkých společností. Článek tím shrnuje klíčové prvky vybraných oblastí Data Governance.

Klíčová slova: Data Governance, Metriky, Data, Datová kvalita, Řídicí proces.

Úvod

Popis problému

S růstem společností a/nebo objemu dat přichází i potřeba data více a efektivně řídit. Pokud není tato potřeba pokryta, může (a zpravidla se tomu tak děje) nastat řada problémů. Nemusí se vždy jednat o nekvalitní data, přestože je to první věc, která vyvstane na mysli.

S absencí, či nedostatečným fungováním, Data Governance však mohou souviset i další problémy, které lze řadit do následujících tří základních kategorií (v závorkách uvádím příklady):

- Finanční (snížený zisk z důvodu neúspěšných kampaní při špatné segmentaci nebo scoringu klientů z důvodu nízké datové kvality)
- Sociologické (zhoršení kultury prostředí na pracovištích a mezilidských vztahů z důvodu absence odpovědné osoby za data a klíčové termíny)
- Technologické (častější "pády" při procesech datové integrace, případně neúspěšné načítání dat do datového skladu)

Zasazení práce do tématického rámce

Tato práce se zabývá následujícími třemi vzájemě úzce souvisejícími oblastmi Data Governance: Metriky, Monitoring metrik, Řídicí proces Data Governance

Metriky

Práce pokrývá výběr klíčových metrik a procesů řízení dat. Z analýzy literatury a vlastní praxe rozlišuje autor dvě základní kategorie metrik Data Governance: Metriky řízení Data Governance a Metriky, které Data Governance využívá pro řízení dat

První kategorie metrik je využívána pro vyhodnocování samotného fungování Data Governance. Druhá kategorie zahrnuje jednotlivé ukazatele, které se využívají pro data. Jinými slovy jednou soustavou řídíme efektivitu Data Governance, druhou pak řídíme data. Vybrané metriky z druhé skupiny mohou figurovat i v první. Často je kladen důraz pouze na metriky datové kvality. Mé chápání, které opírám o

osobní specializaci v rámci studia a více jak tříletou praxi v oblasti Business Intelligence a Data Governance, je širší. Nestačí pouze sledovat data jakožto výsledky práce uživatelů a zpracování aplikačního SW. Je třeba měřit a řídit i procesy a samotné uživatele ideálně ještě před tím, než se chybná data do systémů zanesou. Případně než se zanesou do datového skladu společnosti.

Monitoring metrik Data Governance

Důležitou součástí Data Governance, která umožňuje řídicímu procesu (a odpovědným osobám) dynamicky reagovat na stav dat v podniku a podle něj upravovat pravidla a odpovědnosti, je sledování metrik. Této oblasti se detailně věnuje kapitola Monitoring.

Řídicí proces Data Governance

Popis klíčových prvků procesu řízení Data Governance je věnována pozornost z toho důvodu, že řada společností vnímá Data Governance jako samostatně fungující entitu, která se sama o vše postará ve chvíli, kdy je zavedena. Data Governance však sama za společnost nerozhodne, co dělat, ale může poskytnout směr, kam se ubírat. Na to, aby odpovědné osoby znaly správné postupy, jednali správně a efektivně, je třeba fungující a řízený systém. Jinými slovy, nestačí pouze říci "Máme Data Governance."

Cíle

- Popsat vývoj a aktuální stav DG
- Vybrat "mission critical" metriky DG
- Určit a popsat sledování a proces řízení DG s využitím vybraných metrik

Definice klíčových pojmů

Práce předpokládá základní znalost probírané oblasti, a proto jsou vydefinovány pouze ty nejdůležitější termíny.

K pojmu Data Governance (dále též jako DG) uvádí (SARSFIELD, Steve, 2009) hned několik definic. A takový přístup je správný, protože tyto definice popisují problematiku, kterou se DG sama zabývá - jak nalézt jednotnou verzi pravdy? Různá oddělení, a dokonce i různí zaměstnanci v rámci jednoho oddělení, vnímají tento pojem odlišně. Kniha např. uvádí, že *"Pro top management je data governance o vnesení efektivitu do organizace, zužitkování co nejvíce znalostí ze svých zaměstnanců a zajištění, aby lidé dostali svá data a metriky, které potřebují, aby mohli činit dobrá rozhodnutí"*. Na druhé straně, *"Pro IT je data governance o nastavení platformy pro master data management, chápání, jaká data jsou k dispozici, jaká je třeba čistit a standardizovat, a jak spravovat metadata pro co nejvyšší efektivitu"*. Jinou, dle mého oborově nezávislou, definici uvádí (THE DATA GOVERNANCE INSTITUTE): *"Data Governance je systém pravidel rozhodování a odpovědností pro procesy spojené s informacemi, provozovaný na základě dohodnutých modelů, které definují, kdo může provést dané úkony s danými informacemi, a za jakých podmínek použije dané metody"*.

Data Governance Council (DGC) je nejvyšší autoritou, která definuje všechny klíčové aspekty DG. V zásadě se jedná o zástupce top managementu různých linií organizační struktury (např. ředitel financí, ředitel IT, generální ředitel, ředitel provozu). Z toho vyplývá, že se nejedná o roli na plný úvazek. DGC je svolávána pro řešení nejdůležitějších záležitostí, které mají celopodnikový dopad. Definuje pojmy, pravidla, řeší spory spojené s daty mezi jednotlivými odděleními. V případě menších podniků se lze setkat s velmi omezeným počtem osob připadajících do DGC. V případě, že tuto klíčovou odpovědnost drží jedna osoba, vyskytuje se pojem Head Data Governance Officer. Orientace je zde spíše na business.

Data Governance Office je entita, která analogicky zastupuje v DG střední management - shromažďuje popsané problémy a rozhoduje o jejich prioritizaci. Jedná se zde už o větší až plný úvazek, kde práci vykonávají seniorní datoví Stewardé. Vzhledem k velké odpovědnosti a různorodosti řešených oblastí je třeba, aby pracovníci této skupiny rozuměli velmi dobře jak businessu, tak technologiím.

Data Steward je zjednodušeně řečeno operativní úroveň DG. Stewardi pracují přímo s daty, řeší a komunikují konkrétní situace a problémy. Aktivně zapojují do řešení odpovědné a zainteresované osoby. Orientace zde mírně inklinuje spíše k technologiím. V některých společnostech se Stewardé dělí např. podle

primární odpovědnosti. Existují např. Stewardé, kteří jsou zodpovědní jen za business termíny a pracují s centrálním firemním slovníkem pojmů (business dictionary).

Datová kvalita je velmi široký pojem. Obecně je termín přijímán jako stav dat, kdy jsou vhodná pro daný účel businessu. Jako na metriku se na ní dívá (DATA FOR DEVELOPMENT, INC, 2011): *"Jedná se o metriku, kterou lze měřit hodnotu dat pro podnik"*.

Data Governance v současnosti

Ve velkých společnostech již dnes v řadě případů DG rámec existuje. Nemusí být nutně koncipován do celku s názvem DG. Může existovat po dílčích cílech v rámci jiných governance rámců (celková IT strategie). I přes případy, kdy je DG stanovena, neznamená to nutně vývoj k lepšímu. Problém často totiž tkví ve špatném uplatňování DG pravidel, zaměstnanci z různých oddělení, kterým je přidělena DG role na část úvazku, jsou přetěžováni a třeba ani zcela nerozumí své pozici v té velké mašinérii. DG je živoucí organismus, vyvíjí se, zraje. Odborníci na řízení dat ve společnostech musí pružně reagovat nejen na problémy s daty plynoucí z vnitřku firmy, ale i na tržní prostředí.

Příkladem mohou být finanční instituce, na které spadají stále nové regulační požadavky (Basel III, Solvency II, Fatca, a další), jejichž součástí je mj. i důraz na řízení dat. Pokud je společnost zaměřena pouze na snižování nákladů, snaží se (nejen) v případě Data Governance dosáhnout často lepších výsledků, avšak s méně lidmi a za méně peněz. Bohužel, již v současnosti je počet odborníků na řízení dat v těchto firmách pod potřebnou hranicí, a proto takový postup vede ke spirálovému efektu, kdy dochází k častějším a/nebo větším problémům s daty. Společnosti pak za nižší náklady přicházejí o zisk z ušlých příležitostí, které by měly, kdyby se o svá data staraly. Tak to uvádí (FINANCIAL TIMES BUSINESS LTD., 2012): *"Globální agendou pro finanční instituce se z většiny stává transparentnost. Transparentnost znamená přesná data... aby společnosti splňovaly nařízení Fatca, které přijde v platnost 1. 1. 2013, společnosti budou muset mj. zajistit, aby klientské informace byly centralizovány za účelem snazší identifikace stavu jednotlivých držitelů účtů"*.

S regulačními nařízeními a bezpečností citlivých dat, která bývají zahrnuta do DG, se potýkají i instituce z jiných sektorů. Např. Bostonská Univerzita za tímto účelem inovovala své informační prostředí implementací nového nástroje pro řízení bezpečnosti dat od společnosti Varonis Systems, Inc. Tak uvádí v (BOSTON UNIVERSITY, 2013): *"Univerzita zjistila, že většina současných nástrojů pro prevenci ztráty dat řeší pouze jeden aspekt - kde jsou citlivá data chráněna, ale neviduje již kdo k nim měl po dané období přístup, případně kdo s daty pracoval."*

Metriky pro objekty řízení DG

Tato kapitola má za cíl definovat a popsat klíčové metriky pro vybrané objekty, které se používají v Data Governance.

Metriky pro Data

Data a jejich kvalita jsou dnes stále více vnímány jako aktiva společnosti. Datovou kvalitou se autor zabýval již v článku (VACEK, Martin, 2013), na který tato kapitola navazuje. Jednotlivé dimenze jsou popsány pouze velmi stručně.

Datové metriky patří mezi klíčové především v reaktivním přístupu k řešení problémů s daty. Datovou kvalitu je třeba chápat jako pojem, který lze dále rozdělit na její jednotlivé aspekty, nebo také dimenze. Rozdělení do dimenzí pak pomůže definovat jednotlivá pravidla a metriky pro její měření a cílování aktivit za účelem jejího zvýšení. Na základě výše uvedeného článku jsou zde tedy použity nejčastěji používané dimenze, k nimž jsou navíc přiloženy příklady možných metrik v obecné a konkrétní podobě. Záměrně jsou v příkladech použity různé úrovně detailu. Granularita je řešena v této práci v kapitole věnující se monitoringu:

A. Přesnost

Přesnost vypovídá o tom, jak správně data odráží skutečnost. V případě problémů s přesností mohou tyto problémy mít syntaktický (např. překlepy), nebo semantický charakter (vyplnění hodnoty do špatného sloupce).

Obecný příklad metriky:

- Počet hodnot neodpovídajících těm slovníkovým.

Konkrétní příklad výsledku měření může vypadat takto:

- V databázi aps_clients, tabulce Client, se ve sloupci First_Name vyskytuje 26 % hodnot, které neodpovídají žádné položce v kontrolním seznamu.

B. Úplnost

Úplnost určuje míru vyplnění jednotlivých atributů. V praxi již není zcela běžné rozlišování pouze na povinné a nepovinné hodnoty. V některých případech se pracuje s tzv. "nice to have" atributy. Technicky je prázdná hodnota dovolena, business ji však považuje za důležitou. Při analýze úplnosti se tak mimo NULL hodnot do záporné vyplněnosti mohou zahrnovat právě i zmíněné zástupné hodnoty.

Obecný příklad metriky:

- Výskyt Null (nevyplněných) hodnot u nice-to-have atributů.

Konkrétní příklad výsledku měření může vypadat takto:

- 30 % jedinečných klientů nemá vyplněný telefonní kontakt.

C. Aktuálnost

Aktuálnost dat je dána časem, za jaký se změna informací o sledovaném objektu odrazí v datech.

Obecný příklad metriky:

- Délka trvání propsání změny mezi dvěma systémy.

Konkrétní příklad výsledku měření může vypadat takto:

- Doba trvání změny kontaktních údajů klienta v CRM systému po jejich nahlášení call centru klientem je v 15 % případů delší jak týden.

D. Konzistence

Konzistence vypovídá o různorodosti obsahu dat o jednom unikátním objektu napříč organizací. Ideální stav nastává, když existuje pouze jedna verze pravdy.

Obecný příklad metriky:

- Počet nekonzistentních verzí informací o jedinečném objektu napříč systémy.

Konkrétní příklad výsledku měření může vypadat takto:

- O panu Novákovi (+ např. uvedená část rodného čísla) existují různé informace o trvalém bydlišti v systémech X a Y.

E. Včasnost

Tato časově orientovaná dimenze sleduje, zda uživatelé svá data mají včas pro to, aby na jejich základě mohli provádět správná a včasná rozhodnutí.

Obecný příklad metriky:

- Počet případů, kdy obchodníci neobdrží výsledky data miningu o pravděpodobnosti nákupu včas.

Konkrétní příklad výsledku měření může vypadat takto:

- 30 % informací o pravděpodobnosti nákupu produktu získají příslušní obchodníci s více jak měsíčním zpožděním.

F. Nestálost

Jedná se o frekvenci změn dat v čase. Jedná se spíše o podpůrnou dimenzi, která však může indikovat důležité problémy a může figurovat i v případě tzv. freud monitoringu.

Obecný příklad metriky:

- Počet změn hodnoty atributu za vybrané časové období.

Konkrétní příklad výsledku měření může vypadat takto:

- Informace o původci škody v pojistné události č. XY se za poslední měsíc změnila čtyřikrát.

Metriky pro Procesy

Metriky mohou být navázány na specifické činnosti procesů, které souvisí se vznikem, úpravami a přenosem dat. Spadají sem především časově orientované metriky a lze sem zařadit jak činnosti ryze automatizované (ETL transformace), tak neautomatizované (dodání vytištěných reportů uživatelům).

Pro některé procesy lze s menšími úpravami využít metriky dat definované výše. Např. včasnost dat lze aplikovat přímo na proces kontrolingu. Pokud se zde aplikuje vrstevnost metrik, o které pojednává kapitola Určování úrovně rozsahu a detailu sledované metriky, může pak metrika a její rozpad vypadat následovně:

Nejvyšší úroveň řeší (businessová definice): Má oddělení kontrolingu data včas pro provedení měsíční závěrky do pátého pracovního dne v měsíci?

Data Governance Office pak tuto metriku sleduje na nižší úrovni agregace s více konkrétními omezeními. Ve zmíněném případě tím ví, že je třeba všechna data předat kontrolingu do druhého dne v měsíci, aby byl čas vypracovat závěrku. Dva dny jsou maximální doba trvání. Je třeba počítat s případnou reklamací a provedením potřebných úprav. To je zaneseno již do požadavků na data a průběh procesů.

- Jsou data z jednotlivých procesů předána oblasti kontrolingu do konce prvního dne v měsíci? Případně, kolik procesů zapříčinilo zpoždění dat kontrolingu o více jak den minulý měsíc? Kolik o více, než dva?
- Kolik bylo minulý měsíc evidováno reklamací na dodaná data? Jaký je vývoj hlášených reklamací proti minulým obdobím?
- Kolik reklamací se podařilo vyřešit včas? Jaký je vývoj (ne-)úspěšně vyřešených reklamací včas proti minulým obdobím?

Operativní úroveň řízení dat, která pracuje přímo s daty na základě reportovaných hodnot, vypracovává ve spolupráci s vlastníky dat a vyššími instancemi DG požadavky na změny a některé z nich i provádí. Metriky jsou zde sledovány v nejnižší granularitě, po specifických systémech, pozicích. Příkladem takových metrik může být:

- Jak dlouho trvá datový export ze systému X v procesní činnosti Y?

Metriky pro Role

Role a lidé hrají důležitou roli, pokud jde o řízení dat v organizaci. V (GRIFFIN, Jane, 2011) se uvádí, že jedním ze stěžejních pilířů dobře fungujícího programu Data Governance je odpovědnost. Způsobem, jak tento pilíř postavit pevný, je zajistit, aby příslušní zaměstnanci odpovědnost přijali a zároveň, aby byli odměňováni, pokud překonají své cíle. K tomu je zapotřebí nastavit odpovídajícím způsobem metriky pro jednotlivé role.

Pro nejvyšší úroveň řízení Data Governance (DG Council) je metrikou vyhodnocování finančních přínosů programu DG (viz kapitola Finance DG) a plnění jeho strategických cílů. Metriky dat a procesů se mohou stát přímo KPIs jednotlivých členů DG kompetenčního týmu. Z pohledu úrovně detailu zde platí analogie, tedy pro roli DG Officera se určuje agregovaná metrika oproti metrikám přiřazeným juniornějším členům (Stewardům).

V případě specializace jednotlivých Stewardů je třeba specializovat a vhodným způsobem upravit i metriky k nim přiřazené. V hypotetickém případě může být Datový Steward odpovědný mj. za konkrétní dimenzi datové kvality (např. syntaktickou přesnost). Stav dat z pohledu této dimenze v kombinaci se zavedením korporátního performance managementu se pak stává KPI pro daného zaměstnance.

V jistých případech se může jednat i o týmové metriky, které mají dopad např. na variabilní složku platu jednotlivých pracovníků (SARFIELD, Steve, 2009) uvádí několik příkladů takových metrik převrácených do cílů. Zde je uveden jeden z nich: *"Během šesti měsíců od uvedeného data snižte náklady spojené se zasíláním zboží na neúplné, nebo špatné adresy o 2 % z celkového objemu poštovního"*. Takový cíl je správně definován, protože je měřitelný, obsahuje datum i žádoucí výsledek usilí týmu.

Metriky řízení DG

Prakticky kterákoliv z výše uvedených metrik může figurovat v řídicích metrikách DG. Podmínkou je však její businessový soulad a soulad s cíli nejen DG, ale i požadavky klíčových zástupců businessu.

Jako další příklady lze uvést metriky DG dle jednotlivých oblastí:

- Procesy: délka trvání jednotlivých procesů, počet zaměstnanců zapojených do procesu.
- Organizační struktura: pracnost vyjádřená v člověko-dnech jednotlivých procesů, počet datových Stewardů.
- Nástroje: dostupnost nástroje, počet jedinečných přístupů v čase.
- Obsah: počet požadavků na opravy datových objektů (atributů).

Monitoring metrik DG

Možnosti měření metrik

Některé metriky lze sledovat lidskými zdroji manuálně. Jedná se však o poměrově menší část, než v případě automatizovaného sledování, kdy jsou nad datovými zdroji spouštěny analytické dotazy, případně je na zdroje napojen sofistikovaný softwarový nástroj. Příkladem takového nástroje může být Data Quality Center společnosti Ataccama.

Časové měřítko

Metriky DG lze z pohledu času sledovat dle dvou dílčích hledisek.

Z hlediska průběhu času: v reálném čase a periodicky (denně, měsíčně apod.)

Z hlediska informační potřeby: kontinuálně (bez časového omezení), krátkodobě (ve vymezeném časovém intervalu) a na základě spouštěče, tzv. triggeru (např. v podobě události, kdy je do datového zdroje dokončeno nahrávání dávky dat z jiného úložiště)

Určování úrovně rozsahu a detailu sledované metriky

V praxi je často využíváno vrstvení (agregace) metrik. Pro každou úroveň managementu je sledována jiná úroveň detailu. Top management může zajímat např. "dovolatelnost telefonního čísla", což je velmi obecná metrika. Je proto třeba takové metriky rozdělit do dílčích. Na úrovni Data Governance Office je pak takováto metrika rozložena do většího detailu, který se dále rozpadá v případě operativní úrovně. Datoví Stewardi pak sledují data po jednotlivých dimenzích na úrovni atributů v konkrétních datových zdrojích.

Forma a obsah reportingu DG metrik

Důležitým faktorem reportů je jejich snadná čitelnost a přehlednost. V současnosti je snaha počty reportů omezit, a proto se v DG používají tzv. DG dashboards. Dashboard je zpravidla grafický report, který obsahuje klíčové výsledky měření pro jeho uživatele a je navržen tak, aby se v tisknutelné podobě vešel na jeden list A4. Dashboard obsahuje některé sumární metriky (vyplněnost datových a business slovníků) a detailní metriky (Počty oprav dat, průměrná doba na vyřešení problémů s datovou kvalitou, počet problémů s datovou kvalitou) a jejich porovnání s předešlým obdobím. Zmíněné metriky a reporty budou zajímat spíše vyšší instance DG (tedy DG office a především DG Council). Nižší úrovně DG pak zajímají podrobnější informace. Je pro ně proto navržen jiný dashboard, který sleduje data po jednotlivých dimenzích, v detailu pak nejvíce se vyskytující problémy a nálezy)

Procesy řízení DG

Definice výchozího bodu (existující vs. neexistující DG ve společnosti)

Při tvorbě Data Governance v organizaci je třeba uvažovat, zda již některé procesy (či jiné prvky) již v podniku existují. Lze se setkat hlavně s případy, kdy v podniku existují různě vyzrálé procesy řízení datové kvality.

Pokud jde o velké korporace s již zaběhlou DG praxí, kde se však top management rozhodl o zásadní změny v jejím fungování, je třeba provést analýzu (může se jednat o tzv. "úvodní studii", nebo také "feasibility study"), která zjistí, zda je třeba provést změny ve vybraných oblastech (tzv. revitalizaci), nebo je třeba začít zcela od začátku, protože stávající DG je z většiny nevyhovující.

Mise a klíčové cíle Data Governance by měly odrážet a souviset s celopodnikovou a informační strategií.

Data Governance a některé její klíčové oblasti jsou z části pokryty i v pětiúrovňovém modelu Data Warehousing Maturity Levels, který je uveden v (ARUN, Sen et al., 2012). Z pohledu DG na první úrovni zralosti datového skladu neexistuje žádný jasně definovaný koncept řízení. Na druhé jsou definovány základní role a obecné požadavky na kontroly datové kvality. Třetí úroveň určuje vlastníky klíčových dat, definuje základní rámec DG a prostředky a způsoby formální komunikace. Na čtvrté úrovni se sledují jednotlivé problémy, řízena je datová kvalita spolu s metadaty. Na páté úrovni jsou definovány postupy pro data a metadata change management a aplikují se proaktivní opatření.

Řídící proces DG

Většina zdrojů, které sloužily jako podklad k tomuto článku jako např. (SARSFIELD, Steve, 2009), (ARUN, Sen et al., 2012) a (THE DATA GOVERNANCE INSTITUTE) chápou jako hlavní řídicí proces DG soustavu aktivit, kterou můžeme nazvat Kontinuální zlepšování (angl. Continual Improvement), s kterým se lze setkat např. i v ITIL v3. Jeho zjednodušenou strukturu uvádím v Obr. 1.



Obr. 1 - Zjednodušený koncept Kontinuálního zlepšení

Zdroj: Autor

Jak je uvedeno v kap. Monitoring, reportování nálezu může být automatizované (výsledek dotazu nad databází), nebo manuální. Do manuálního spadá např. i stížnost. I v takové případy mohou ve finále iniciovat změnové řízení na samotnou DG.

K upřesnění: ne každý nález/problém musí nutně znamenat iniciaci změnového řízení. Řešení některých z nich může být nákladných a potenciální přínos jejich vyřešení naopak nízký. V druhém kroku již může být nález uznán za žádoucí k řešení. V tom případě jediné, co se provede, je zaevidování tohoto nálezu. Tento proces probíhá na úrovni DG Office. Pokud je nález významný a/nebo může znamenat jeho výskyt, nebo řešení, značné finanční náklady, je definovaný detailní požadavek předán vyšší autoritě (úroveň DG Council), která uskuteční v pořadí druhou validaci nálezu a dopadů (ne-)řešení.

Změnové řízení se může týkat zásahů do informačních systémů, změny definic pojmů, ale i změny přiřazení rolí, jejich pravomocí, úpravy prahových hodnot sledovaných metrik a dalších a to jednotlivě i v kombinaci.

Finance DG

Řídící proces DG vychází z definice jejích cílů. Primárním motivačním faktorem managementu v komerční sféře je zisk, a proto existuje tendence vyjadřovat úspěšnost objemem finančních prostředků, které dané řešení má přinést (nebo přineslo), a to jak přímo, tak nepřímým způsobem. Důležité je, aby cíle DG nezůstaly pouze na nejvyšší úrovni vedení. Zavedení DG do společnosti znamená i jisté kulturní změny, a proto je třeba zaměstnancům cíle a plánované přínosy DG jasně komunikovat, což má za úkol nejvyšší orgán - Data Governance Council (GRIFFIN, Jane, 2011).

Financemi DG se zabývá diplomová práce (KMOCH, Václav, 2011), která uvádí následující přístupy ke sledování financí:

- ROI (Return on Investment; Návrstnost investic) - tento ukazatel se počítá jako poměr výnosů a investic převedený do procent. V práci je uveden i upravený vzorec:
- $ROI [\%] = \text{výnosy} * \text{míra investice DG na výnosu} / \text{investice DG} * 100$

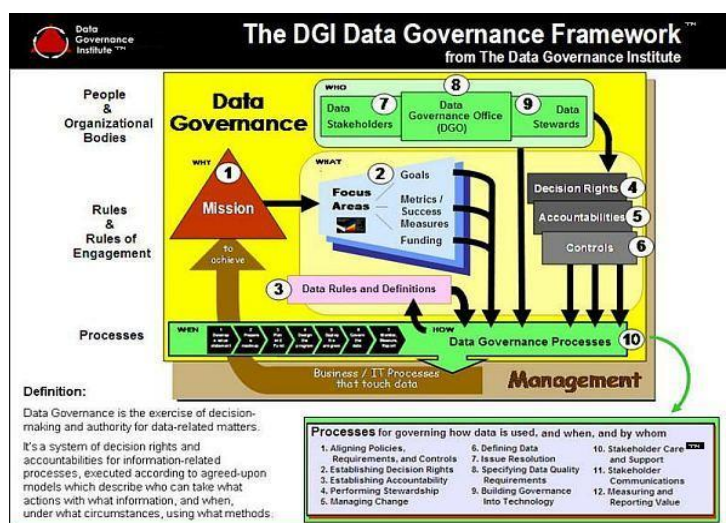
- CIDDA (Confidence in Data-Dependent Assumption; míra jistoty datově závislých odhadů). Jedná se o číselný ukazatel v rozmezí 0 až 1 vyjadřující míru jistoty, kterou lze dále výpočet ROI ovlivnit.

$$\text{CIDDA} = G * M * TS$$

- G je míra jistoty o dobré datové kvalitě (je třeba vymezit pojem „dobrá“ kvalita dat)
- M je míra jistoty o sémantice dat
- TS je míra jistoty, že data jsou důvěryhodná

Vybrané procesy DG

Procesy můžeme rozdělit na nastavovací "setup" procesy (např. nastavení cílů, rolí, odpovědností a dílčích procesů), komunikační procesy (monitoring, reporting, validace) a change management procesy (řešení problémů a nálezů). Do těchto kategorií lze zařadit 12 klíčových procesních oblastí, které uvádí Obr. 2.



Obr. 2 - DGI Data Governance Framework

Zdroj: <http://www.datagovernance.com>

Před tím, než je zahájena práce s definováním procesů řízení dat, je třeba si nejprve odpovědět na několik klíčových otázek, které určí procesní rámec, tedy co bude DG pokryto, a úroveň vyspělosti jeho jednotlivých procesů.

Zajímavý pohled na tuto problematiku uvádí (POWER, Dan, 2011): *"Procesy vyžadují přemýšlet o tom, co data governance bude dělat ve chvíli, kdy je zavedena. Jaká je její mise a řád? Které firemní systémy bude sledovat z pohledu dat? Jak budou stanoveny pravidla a jakým způsobem budou prosazována? Jak budou měřeny pozitivní dopady těchto pravidel?"*

Procesy DG jsou v zásadě definovány napříč společnostmi, tzn. že nejsou definovány individuální DG procesy pro jednotlivé systémy, protože v tom případě by byla administrativa spojená s řízením dat neefektivní - v řadě případů je totiž třeba řešit problémy ve více systémech, nebo v rámci datových toků mezi nimi.

Specifika DG v podnicích různého typu

Na podniky se můžeme dívat z pohledu sektoru (finanční, zdravotní, těžařský, státní,...), nebo z pohledu velikosti (korporace, malé a střední podniky).

V prvním případě lze správně předpokládat, že datový obsah napříč sektory bude různorodý. Dle mého je však potřeba řídit data, včetně přístupů potřebných pro zavedení fungující DG, z většiny stejná viz uvedení podobností potřeb pro řízení bezpečnosti a auditovatelnosti organizací v kapitole Data Governance v současnosti.

Finanční a vzdělávací sektory v této práci byly již zmíněny. Jako další důkaz, který podporuje toto tvrzení lze zdravotnický sektor. V zemích po celém světě jsou data o pacientech velmi decentralizována.

Nemocnice a kliniky jsou držitelé obrovského množství dat o svých pacientech. Někde se zavádějí řešení jako patientské e-karty (podobný projekt již byl i naší vládě předložen). Informace o pacientech jsou velmi citlivými údaji a je třeba splňovat řadu důležitých zákonů a nariadení, aby byla jejich data chráněna proti zneužití. Postoj zdravotnického sektoru k DG popisuje (REEVES, Mary G and Bowen, Rita, 2013): *"S příchodem elektronických záznamů o pacientech a digitálních zdravotních informací, data v elektronických zdravotních záznamech (EHR z angl. electronic health records) nyní představují platný zdravotní záznam. Odpovědní pracovníci za tyto záznamy musí nyní pracovat ruku v ruce s IT a systémovými experty, aby tak reagovali na nové výzvy, maximalizovali příležitosti a snižovali rizika"*. Existují samozřejmě sektorová specifika. Článek se zmiňuje o tzv. CMIO (Chief Medical Information Officer) jakožto členovi top managementu, který se stává i členem nejvyššího orgánu DG.

Pokud budeme na podniky nahlížet jako na různě velké celky, zde se potřeby liší. Existují volně dostupné i placené frameworky a komplexní konzultantské služby, které nabízejí zavedení Data Governance do organizace. Nedostatek však lze nalézt škálovatelnosti takových řešení, protože většina z nich počítá, že uživatelem (klientem) DG bude velká korporace. Postoj SME sektoru (Small and medium enterprises - malé a střední podniky) k takovým řešením je většinou negativní, protože taková naddimenzovaná řešení znamenají obrovskou finanční zátěž. Nejen úroveň detailu, ale i pokrytí konkrétních datových oblastí, řeší tyto podniky.

(BEGG, Carolyn and Caira, Tom, 2011) se problematice DG v SME podrobně věnuje: *"V případě SME se společnosti podobné velikosti zaměřují na různé požadavky na přístup a používání dat. Společnost A nevyužívá vůbec e-business jinak než prostřednictvím e-mailů, a proto se hlavně zaměřuje na přístup k nim pouze interně. Společnost B investovala do e-business facilit, a tedy využívá webových služeb jak pro zákaznické, tak dodavatelské vztahy. Takto musí pak společnost B řešit zaměření na data nejen interně, ale i externě."* Jsou zde zmíněny jak aspekty škálovatelnosti, tak různorodosti zaměření na konkrétní datové oblasti. V malých podnicích, narozdíl od těch velkých, jsou datové oblasti buď řešeny, nebo neřešeny vůbec. Vhodným by se pak jevil škálovatelný modulový DG rámec v podobě jednoduchých obecných doporučení. V praxi se jim někdy též říká HLC (High Level Concept).

Závěr

Práce řeší sledování kvality a řízení dat. Identifikuje klíčové zástupce ze dvou oblastí data governance - procesů a metrik. Zaměření je nejen na metriky spojené s datovou kvalitou, ale i na metriky v rámci společnosti, které přispívají k vyšší efektivitě řízení dat včetně těch, kterými se hodnotí samotný program (iniciativa) data governance. Z práce vyplývá, že metriky přiřazované procesům a rolím mohou být upravenou verzí metrik samotných dat.

Data Governance není jen o měření datové kvality, ale i o měření a řízení dalších objektů působících v rámci společnosti, které korporátní data ovlivňují. Na práci lze navázat rozpracováním dalších předmětných domén, z kterých se DG skládá, jako např. role a nástroje.

Literatura

ARUN, Sen, K (Ram) RAMAMURTHY, and Atish P SINHA. 2012. A Model of Data Warehousing Proces Maturity. *IEEE Transactions on Software Engineering*. **38**(2).

BEGG, Carolyn and Tom CAIRA. 2011. Data Governance in Practice: The SME Quandary Reflections on the Reality of Data Governance in the Small to Medium Enterprise (SME) Sector. *European Conference on Information Management and Evaluation*.

BOSTON UNIVERSITY. 2013. Boston University Leverages Varonis' Data Governance Solution. *Professional Services Close - Up*.

DATA FOR DEVELOPMENT, INC. 2011. *What is Data Quality?*. [online]. [Accessed duben 2013]. Available from World Wide Web: <"http://www.dfdi.com/whatisdq.htm" >

FINANCIAL TIMES BUSINESS LTD. 2012. Corporate statement: Smartstream - Best practice in data governance. *The Banker*.

GRIFFIN, Jane. 2011. Data Governance Defined Power to the People: Build A Culture Of Accountability That Rests On Foundational Pillars Of Education, Buy-in, Responsibility and Communication. *Information Management*. **14**(6).

KMOCH, Václav. 2011. *Data Governance - koncept projektu zavedení procesu*. Vysoká Škola Ekonomická.

POWER, Dan. 2011. A 4-D Approach to Data Governances: Work across peopl, process, technology - and information - to achieve data governance sucess. *Information Management*. **14**(6).

REEVES, Mary G and Rita BOWEN. 2013. Developing a data governance model in health care. *Healthcare Financial Management*. **67**(2).

SARFIELD, Steve. 2009. *The Data Governance Imperative*. IT Governance Publishing.

THE DATA GOVERNANCE INSTITUTE. *Definitions of Data Governance*. [online]. [Accessed 28 Dec 2013]. Available from World Wide Web: "http://www.datagovernance.com/adg_data_governance_definition.html" >

VACEK, Martin. 2013. Dimenze datové kvality a nástroje a metody pro její zlepšování v podniku. *Systémová integrace*. **20**(3).

JEL Classification: O30

Summary

Metrics, monitoring and control process of Data Governance

This article identifies and describes important subject areas of Data Governance - metrics and processes with emphasis on control processes. Based on analysis of Czech and foreign literature and practical experience of the author an information framework is set including definitions of key words like Data Governance itself, Data quality, Data Governance Council. There is also description of current status of Data Governance in the world. Specific challenges the data governance specialists face are mentioned such as low trust of stakeholders in return on investment of Data Governance and new regulatory requirements. These problems are faced by organizations across all economic sectors. Following are sets of metrics used for governing data grouped by individual Data Governance objects (data, roles, processes). Next chapter is focused on description of monitoring and reporting of mentioned metrics. Monitoring is described mainly from time and automation views. Current trend of reporting is mentioned as it's being focused on providing information about data to responsible persons effectively by using so called Data Governance Dashboards. Related to this is another sub-chapter that focuses on setting level of detail for specific Data Governance entities and different levels of management. There is also a part dedicated to financial management where some benefits calculation approaches are mentioned. Another part describes common characteristics as well as differences between organizations of various economic sectors and sizes. This is the way the article summarizes key concepts and elements of chosen Data Governance areas.

Key words: Data Governance, control process, metrics, data quality.

Příprava modelu multimediální komunikace

Zdeněk Vondra

zdenek.vondra@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: doc. Ing. Stanislav Horný, CSc., (horny@vse.cz)

Abstrakt: Článek se zabývá aktuálním stavem příprav modelu multimediální komunikace zahájených v roce 2011 v souvislosti se založením Grafické a multimediální laboratoře Fakulty informatiky a statistiky VŠE. Požadavek na vznik modelu vychází z nedostatku konceptuálních metodických podkladů pro práci s multimediálními prostředky bez přímé návaznosti na jeden konkrétní nástroj, které by popsali komunikaci jako celek. Existují obecná pravidla, která zatím popisují jednotlivé konkrétní multidisciplinární úkazy nebo problémy.

V důsledku digitalizace jsou dostupné téměř všechny multimediální nástroje široké veřejnosti a tak se stává klíčovým nástrojem jejich využití právě silná konceptuální úroveň, která je využitelná ve více oblastech multimédií najednou. V praxi to znamená, že například u reklamních agentur dochází k navyšování hodnoty tvůrčího konceptu nad vlastní výrobou mediálního obsahu, který je díky tomu mnohdy outsourcován.

Přípravy modelu multimediální komunikace vychází z tradičních modelů a teorií. V práci je uvažována skutečnost, že komunikace představuje mezioborovou oblast a tak je k ní také přistupováno. V článku jsou shrnuta aktuální východiska, metody a postup tvorby modelu. Jsou prezentovány dvě úrovně ověřování a to nástrojová a společenská, která je odvozena od různých oblastí využívající multimediální nástroje. Ve třetí třetině jsou prezentovány nové poznatky, které posouvají model dále ve vývoji a jsou formulovány další úkoly pro práci na modelu.

Klíčová slova: multimédia, multimediální komunikace, multimediální design, model komunikace, duální kódování, interaktivita, produkční řetězec.

Úvod

O řešení problému, účel práce a očekávání

Jako student doktorského studia aplikované informatiky se v rámci tématu své disertační práce zabývám formulací obecného modelu multimediální komunikace. Účelem práce na modelu a jeho očekávaným přínosem je poskytnutí rámce pro zvýšení efektivity tvorby multimediálních výstupů a tím i zlepšení komunikace, která tyto výstupy používá. Očekávání od nového modelu jsou definice obecných zásad platných napříč oblastmi multimédií, které usnadní jejich porozumění a tvorbu ve vzájemném kontextu a synergii. Existence modelu multimediální komunikace poskytne platformu pro optimalizaci a tvorbu informačních a komunikačních systémů, které budou schopny vlastní komunikační, ale především výrobní procesy multimediálních prostředků podporovat a to především v procesní, projektové i ekonomické rovině. V tomto ohledu je vhodné nahlížet na komunikační model dvojím způsobem a to jednak jako na model samotný a za druhé jako na sdružený model výrobního (produkčního) řetězce multimediálních výstupů (v textu bude vysvětleno).

Položená otázka na hledání obecného modelu multimediální komunikace byla iniciována vlastními zkušenostmi s tvorbou počítačové grafiky, digitálního videa a zvuku. V souvislosti se studiem informačního managementu a přístupů postavených na Checklandově metodologii měkkých systémů jsem dospěl k hypotéze, že pokud člověk konceptuálně porozumí jedné z těchto oblastí, je schopen abstrahovat tyto principy (oprostit se od nástrojů) a dále je přenášet, umožní mu to jednoduší vhled a vstup do dalších konceptuálně příbuzných oblastí. Nároky na jejich pochopení budou díky tomu z větší části na úrovni osvojení práce s tvůrčími nástroji než na úrovni pochopení principu výstavby komunikace. Osobně jsem registroval tento pocit při přechodu z tvorby počítačové grafiky na tvorbu digitálního videa.

Používání multimediálních nástrojů v kontextu digitalizace

Člověka obklopuje obrovské množství komunikací, z nichž poměrně významná část nedokáže kvalitně přenést sdělení. Odkud pramení tato neefektivita? Marshall McLuhan předpověděl, že média podobně

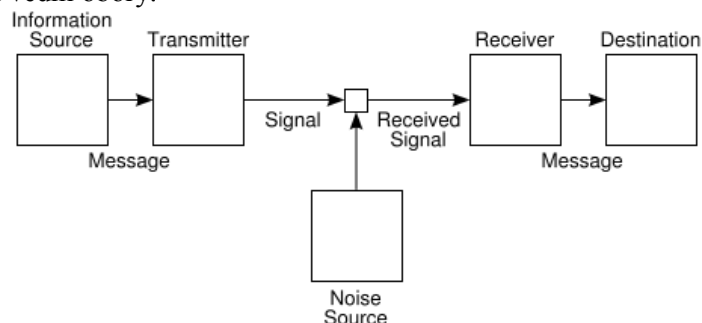
jako jiné technologicky dotčené oblasti se budou nástrojově vyvíjet tak rychle, že všeobecná lidská znalost pro jejich užívání nebude tomuto trendu stačit [McLuhan, 2011]. Základní průlom pro oblast médií a multimédií byla digitalizace, která značně snížila výrobní a pořizovací náklady, čímž masově rozšířila možnosti tvorby a využití mediálních a multimediálních nástrojů. Díky dostupnosti technologií se postupně utvořila velmi početná skupina polo-profesionálů a laiků, kteří pouhým vlastněním některého z nástrojů pro tvorbu mají možnost tvořit komunikace pro široké publikum (zejména skrze internet). Každé médium nebo multimédium bylo v období před digitalizací velmi technologicky a postupově náročné, což činilo nároky na vytvoření úzce profilovaných skupin profesionálů ovládajících složité konceptuální i nástrojové postupy. Digitalizace, která slučuje výrobu na architekturu výpočetní techniky, vytvořila budoucí poptávku po profesionálech šířeji orientovaných, avšak vždy se znalostí ovládání výpočetní techniky a schopností přenášet vzniklá data mezi jednotlivými platformami. Multi-disciplinární přístup zároveň vyprovokoval další pokrok, který standardní nabídku mediálních nástrojů (médiatypů) rozšířil o nové formy. Do popředí se tak dostává konceptuální rovina, která v okamžiku obecné dostupnosti všech komunikačních prostředků představuje rozdílovou veličinu (konkurenční výhodu).

Model multimediální komunikace by měl přispět v tomto okamžiku jako nástroj pochopení konceptuální roviny, který umožní rychleji přenášet koncept z jednoho typu nástroje do druhého nebo extrakce různých typů komunikací z jednoho původního záměru. Lidé s tvůrčím nadáním nebo iniciativou zpravidla nejsou ochotni věnovat mnoho času analytické práci, která by jejich snažení nasměrovala. Model má za jeden ze svých cílů usnadnit a přiblížit tento analytický přístup, aby nečinil překážku a díky tomu nebyl vynecháván.

Komunikace a inspirativní modely

Na komunikace obecně se snažím nahlížet analyticky jako na proces zabývající se přenosem informace za účelem požadované reakce nebo vytvoření znalosti. Multimediální komunikace je každá taková, která využívá multimediálních výstupů za tímto účelem. Tento přístup odpovídá oblasti teoretické a aplikované informatiky. Při studiu komunikace je nutné vycházet z předpokladu, že se jedná o zásadní proces existence sociálních i technických systémů, a že přístupů k jeho pochopení a vysvětlení je více a že se vzájemně ovlivňují (například rétorický, sémiotický, fenomenologický, kybernetický, socio-psychologický, kritický, socio-kulturní dle [Craig, 1999]). Pro další práci je kladen důraz na ty poznatky, které blíže souvisí s efektivitou přenosu sdělení skrze multimediální nástroje.

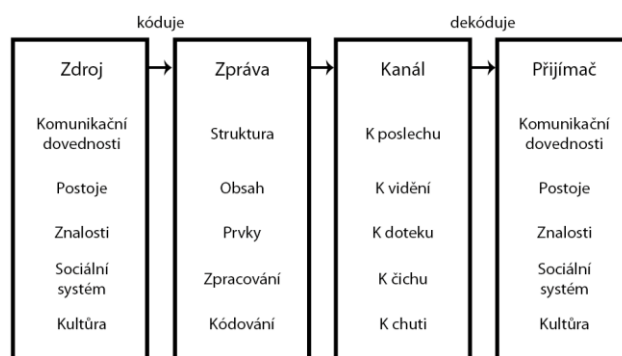
Inspirací pro tvorbu samostatného modelu multimediální komunikace byl Shannonův-Weaverův model poprvé prezentovaný v rámci matematické teorie komunikace v roce 1948 (obrázek 1). Představoval nástroj důležitého pokroku v chápání komunikace ve strojích, který měl přesah i do jiných mnohdy i neexaktních oborů. V rámci studia informatiky nám byl tento model prezentován jako základní odrazový můstek pro chápání procesu komunikace. Shannonův-Weaverův model komunikace byl inspirací pro další modely, které postupně upřesňují jednotlivá témata procesu komunikace v návaznosti na konkrétní přístupy a související vědní obory.



Obrázek 17: Shannon-Weaverův model komunikace

Zdroj: Shannon, 1948.

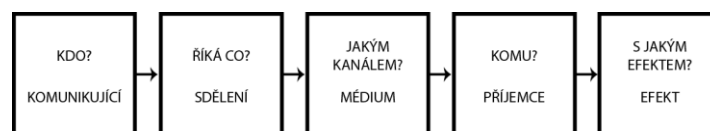
Příklad navazujícího modelu, který se snaží o rozšíření jeho platnosti, představil Berlo v roce 1960 (obrázek 2). Byl to Source-Message-Channel-Receiver model (zkráceně SMCR), který navazuje na Shannon-Weaverův a aplikuje jej na mezilidskou komunikaci, čemuž odpovídají i parametry, které ve svém modelu prezentuje. Model představuje jednosměrný tok informací, který připodobňuje mezilidskou komunikaci telefonnímu spojení [Corman, 2007].



Obrázek 18: Berlův SMCR model komunikace Zdroj: autor dle communicationtheory.org

Uvažovaný model multimediální komunikace využívá obdobně jako Berlův vytváření přesahu na platformě již existujících modelů. Vzhledem k měkké povaze komunikačního procesu, ve kterém je přítomen člověk (na pozici iniciátora, realizátora i recipienta) bylo třeba hledat odrazový můstek i mezi modely primárně humanitně orientovanými.

Základním takovým konceptem je Lasswellův model komunikace (obrázek 3) [Lasswell, 1948]. Lasswell popsal symbolicky také v roce 1948 proces komunikace za účelem analýzy procesu masové komunikace. Každému prvku komunikačního procesu přisoudil otázku a vytvořil prostor pro odpovídající analýzu. Problémem modelu bylo neuvažování zpětné vazby. Na Lasswellův model navazují stejně jako na Shannon-Weaverův také další modely jdoucí až například ke konstruktivistickému modelu nebo transakční komunikaci.



Obrázek 19: Lasswellův model komunikace Zdroj: autor dle communicationtheory.org

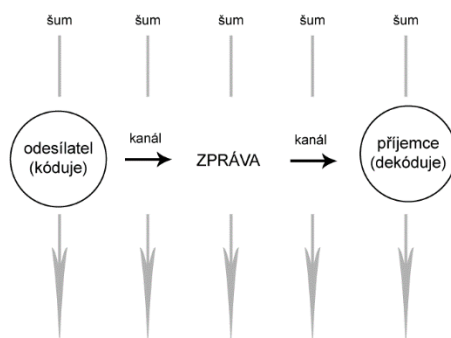
Inspirace Lasswellovým a Shannon-Weaverovým modelem dává prostor pro vzájemné prolínání a využívání různých přístupů, které sledují obdobný nebo stejný cíl.

Schramm prezentoval v padesátých letech další přístup zaměřený a zúžený na dopad sdělení na jeho příjemce. Tato část komunikačního modelu je místem vytváření znalosti. Sdělení působí na recipienta skrze obsah a formu, které dohromady utváří sdělení samotné. Základním obsahem sdělení z hlediska efektu jsou doporučení, povel nebo otázka. Schramm se zaměřil na aspekt komunikačních schopností (gramotnosti) a popis procesů, které participují v přenosu a utváření znalostí. Komunikace je popsána coby sociální interakce, kde alespoň dva účastníci (subjekty) sdílejí znakové sady a sémiotická pravidla. Není tedy uvažována komunikace sám se sebou. V tomto kontextu byly popsány přístupy [Schramm, 1954]:

- Syntaxe (formalizovaný zápis)
- Pragmatika (vztah mezi znaky, výrazy a příjemci)
- Sémantika (vztahy mezi znaky, symboly a tím co reprezentují)

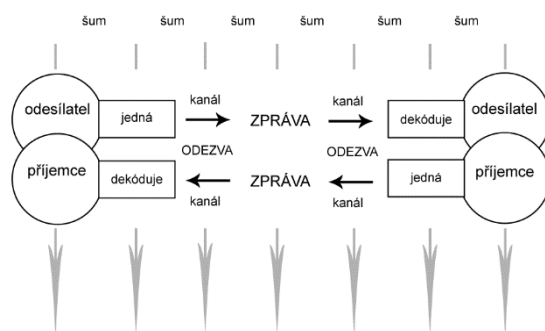
Barnlund představil v roce 1970 transakční model komunikace, který je označován jako konstitutivní model. Model je stále považován za současnou úroveň poznání mezilidské komunikace. Transakční komunikace představuje odklon od klasického jednosměrného lineárního schématu, jak tomu bylo u předcházejících modelů. Počítá se zapojením vysílače komunikace i recipienta coby aktivních prvků a předpokládá jejich interaktivitu. Příjem (dekódování) a vysílání (kódování) sdělení probíhají zároveň, ve dvou směrech a vzájemně se ovlivňují. Komunikace probíhá jako skupina transakcí, které dohromady skládají celé sdělení. To má dopad na vyšší míru porozumění, neboť pokud má například příjemce pocit, že nechápe, nebo vysílač, že jej příjemce špatně chápe, tak komunikaci adekvátním způsobem upraví změnou nebo doplněním chybějících nebo špatně pochopených transakcí (slov, vět, metafor, obrazů). Transakční model chápe komunikaci jako neustálý proces vzájemné regulace za účelem přenosu sdělení a vzájemného pochopení [Barnlund, 2008].

Pro ilustraci časového a věcného přesahu výše uvedených modelů prezentuji další přístup, kterým je Rothwellův výklad komunikace a interaktivity z roku 2009 [Rothwell, 2009], který prezentuje lineární (obrázek 4), interaktivní (obrázek 5) a transakční (obrázek 6) komunikaci ve vzájemném porovnání.



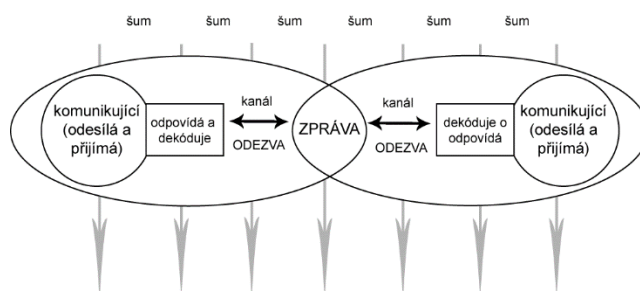
Obrázek 20: Lineární komunikace dle Rothwella

Zdroj: autor dle Rothwell, 2009



Obrázek 21: Interaktivní komunikace dle Rothwella

Zdroj: autor dle Rothwell, 2009



Obrázek 22: Transakční komunikace dle Rothwella

Zdroj: autor dle Rothwell, 2009

Z hlediska přístupu k multimediální komunikaci považují Rothwellovu interaktivní komunikaci za interaktivní v omezené míře, neboť se fakticky jedná o sérii lineárních komunikací s opačnou orientací, které sice značí obousměrnost komunikace avšak s omezením na úrovni časové prodlevy, která může komunikaci a přenos jejího sdělení zásadně posunout. Za jednoznačně a plně interaktivní považují transakční komunikaci. Zpětná vazba byla zpočátku u humanitně orientovaných modelů opomíjena. Důsledkem tlaku na její začlenění je uvažování interaktivity i u humanitně orientovaných modelů, což je i nyní společensky patrné, kdy rozšířená masová komunikace částečně ustupuje komunikaci, která je segmentovaná dle dílčích cílových skupin.

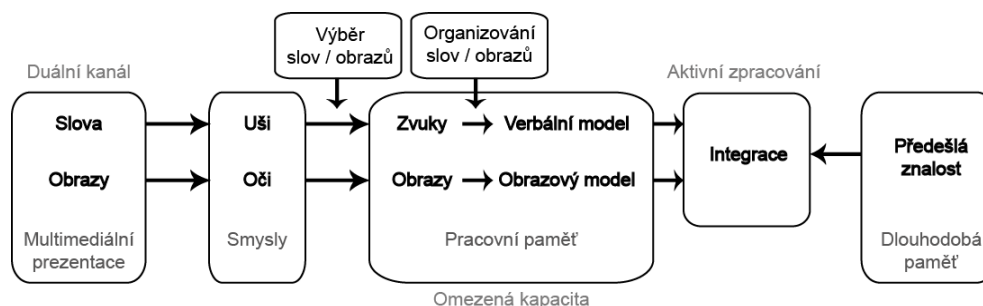
Jaký je přínos studia modelů komunikace?

Modely se shodují v tom, jakou strukturu komunikační proces má a jaké otázky v souvislosti s jednotlivými prvky řeší. Budoucí model multimediální komunikace se tak může opřít o fungující schéma prvků. Z hlediska vlastností a uvažovaných principů poskytují modely pomyslnou paletu charakteristik, z nichž je možné vybrat ty, které se projeví během užívání multimédií. Ověření původní myšlenky však ukázalo, že téměř žádnou z charakteristik nelze opomíjet, neboť každá komunikace sebou nese nástroj, který naplňuje fungování média. Multimediální komunikaci tedy nelze chápat jako formu, kdy multimédia představují něco navíc, ale jako formu, kdy multimédia představují nástroj, který je příslušný každé z forem komunikace coby nástroj kódování a později dekódování. V kontextu digitalizace lze říci, že

multimédia lze částečně a ne úplně přesně vymezit skrze platformu digitálního zpracování textu, fotografií, grafiky, zvuku, animací a videa [Vaughan, 2008].

Multimédia

Multimédia jsou zjednodušeně ty komunikační nástroje, které působí na recipienta pomocí více než jedné sdělovací roviny, kdy by tyto roviny měly být ve vzájemném vztahu a synergii (jedno médium doplňuje druhé a naopak). U recipienta komunikace je tak mířeno sdruženými nástroji nesoucími sdělení na více jeho příjmových rovin současně, čímž je vytvářen intenzivnější efekt komunikace. Klíčovým východiskem vlastní pro studium zásad fungování a efektu multimédií na přenos informací a tvorbu znalostí je práce prof. Mayera, který zkoumá dopad multimédií na proces učení (v rámci pedagogických aktivit). Mayer v rámci svých závěrů formuloval sedm principů pro tvorbu multimediálních výstupů [Mayer, 2009]. Klíčovým principem je vlastní princip multimediality, který je postaven na teorii duálního kódování (obrázek 7), kterou formuloval Paivio [Paivio, 1986]. Tento princip specifikuje, že recipient přijímá obsah komunikace jako obrazové a verbální struktury, které se vzájemně posilují a působením na více vjemových rovin vytvářejí komplexnější obraz sdělení a jednodušeji tak tvoří znalost nebo iniciují reakci. Tato specifika odpovídají senzorickeému (kognitivnímu) přístupu pro výklad multimédií, který zkoumá především psychologie.



Obrázek 23: Princip duálního kódování

Zdroj: autor dle Mayer, 2009

Při studiu multimédií jako nástroje komunikace narážíme na mnoho různých fenoménů a poznatků, které je často obtížné vzájemně vymezit a uspořádat, neboť jsou odvozené od různých přístupů. Oblast multimédií se tak v kontextu výkladu fungování komunikace jeví jako relativně nepřehledná.

Pro oblast multimédií upřesnil Mayer tři specifické dílčí přístupy [Mayer, 2009]:

- **Doručovací přístup** se zaměřuje na technologickou rovinu. Specifikuje postupy fungování technologické architektury procesu komunikace na hardwarové i softwarové úrovni. Pro práci na modelu představuje technologické možnosti a ekonomické limity na zdroje pro realizaci konkrétních komunikací. Tato oblast se v rámci výzkumu a vývoje zaměřuje na optimalizaci stávajících forem komunikace, orientaci ve vytvořeném datovém obsahu a tvorbě nových technologických nástrojů.
- **Prezentační přístup** označuje vlastní postupy ztvárnění obsahu komunikace. Představuje tvůrčí rámec pro výběr prostředků a komunikačních postupů. Definiuje, jaké informace budou jak zobrazované a jaký jim bude vložen smysl. Do této oblasti se odráží nové poznatky z designu a sémiotiky. Nelze opomíjet, že součástí sdělení je kromě obsahu také forma (viz Schramm).
- **Senzorický (kognitivní) přístup**, který byl již výše zmíněn, je postaven na studiu efektu komunikace u jejího příjemce. Cílová skupina (mluvíme-li o komunikaci směřované lidem a ne strojem) představuje primární měkkou složku plánování efektu komunikace (pomineme-li lidské zdroje potřebné k tvorbě komunikace). Pomocí tohoto přístupu jsme částečně schopni odhadovat a částečně i plánovat efekt celé komunikace prostřednictvím kvalifikovaného přizpůsobení prezentovaného obsahu a způsobu jeho doručení. Kognitivní aspekty jsou patrné nejen na straně příjemce, ale také na straně tvůrce, neboť mezi těmito protějšky není nikdy rovnost. Tyto odlišnosti jsou iniciátorem podrobné analytické práce, která se snaží tyto pohledy co nejvíce přiblížit, respektive přiblížit tvůrce příjemci. Proto je jedním ze základních nároků na práci tvůrce velká míra empatie a odhadu lidí, které vedou právě k analytickému přístupu.

V analytické práci a přípravě konkrétní komunikace nelze tyto přístupy od sebe oddělit, ale je vhodné znát jejich hranice, aby byl schopen tvůrce komunikace měnit svůj přístup ke každému dílčímu úkolu na základě znalosti jeho zařazení.

Bližší specifikace metod

Práce na modelu multimediální komunikace probíhá pozorováním a popisem reality v konkrétních oblastech, kdy poté dochází k indukci poznatků a formulaci obecných pravidel. Abstrahované principy jsou vzájemně validovány, upřesňovány a sjednoceny do společné verze a poté znovu už jen jako jeden společný princip implementovány do zkoumaných oblastí. V tuto chvíli dochází k výkladu praxe skrze odvozenou teorii. Pokud výklad alespoň v jedné oblasti neobstojí, je znovu abstrahován a ověřován na obecné úrovni s vědomím nepřesností, které vznikly. Pokud zjištěný princip obstojí, je s ním dále pracováno, čímž jeho ověřování nekončí. Stav, kdy bude možné říci, že model je hotov nelze stanovit, neboť dobu ověřování poznatků nelze jednoznačně určit. Níže budou zmíněny oblasti, ve kterých je model ověřován. Ty by měli být doplněny o další tak, aby záběr a platnost modelu mohla být co nejširší. Lze tedy nyní hovořit spíše o aktuálním modelu multimediální komunikace.

Způsob tvorby a ověřování modelu je aktuálně postaven ve dvou vrstvách. První vrstva představuje zkoumání principů fungování a tvorby jednotlivých nástrojů multimédií prakticky popsanych dle Vaughana (text, fotografie, grafika, zvuk, animace, video) [Vaughan, 2008]. Druhou vrstvu představuje aplikace těchto nástrojů do specifických oblastí společenského významu vzájemně odlišných na základě vnitřních charakteristik motivace pro komunikaci:

- Marketingová komunikace (aplikace a tvorba reklamy)
- Animace kulturního dědictví (komunikace a využití hodnot kulturního dědictví)
- Alternativní divadelní tvorba (představení, výstavy a instalace)

Každá z oblastí vkládá do přípravy obecného modelu nové sobě specifické poznatky, které mohou být následně viditelné i v ostatních oblastech.

Marketingová komunikace popisuje praxi reklamních agentur. Jedná se o společnosti fungující v tržním prostředí, jejichž produktem je tvorba komunikací a cílem dosažení vlastního ekonomického výsledku i ekonomického výsledku iniciátora komunikace (zadavatele). Výzkum v této oblasti je postaven na osobních rozhovorech a kvalifikovaném pozorování. Získané informace jsou sumarizovány, ověřovány a používány pro účely praktické výuky (cvičení kurzu Řízení multimediálních projektů v rámci profesně orientovaného bakalářského oboru Multimédia v ekonomické praxi). Získané informace v rámci příprav výuky prochází konceptuálním a významovým rozkladem, kdy dochází k hledání doplňujících informací a vysvětlujících principů. Získané poznatky jsou postupně generalizovány tak, aby jejich přínos byl využitelný širěji. Studenti se na druhé straně v rámci výuky účastní získávání dat pro tvorbu modelu v rámci rešerší komunikací na trhu, čímž jsou mimo vlastní edukace zapojeni do procesu konceptuálního vysvětlování jednotlivých komunikací. Vzniklý dialog napomáhá získávání nových relevantních poznatků.

Animace kulturního dědictví se zaměřením na jeho komunikaci a prezentaci vychází z vědeckého projektu NAKI, který se zabývá uchováním národní kulturní identity. Grafická a multimediální laboratoř má za svůj dílčí úkol vytvořit metodiku multimediální prezentace kulturního dědictví. V této oblasti jsou základní principy komunikace zkoumány ve vztahu k úkolům animace kulturního dědictví. Povaha úkolu komunikace v této oblasti nestojí na ekonomické výtěžnosti, ale na schopnosti efektivního sdělování informačního obsahu za účelem podpory vědomí české společnosti o vlastním historickém a kulturním zázemí. Jsou pojmenovány a oslovovány různé cílové skupiny za stejným účelem, kdy je zpracováván v kontextu animace především stupeň odbornosti v závislosti na oslovení laické nebo odborné veřejnosti.

Alternativní divadelní tvorba probíhá ve spolupráci Katedry alternativního a loutkového divadla DAMU, Katedry arts managementu FPH VŠE a Grafické a multimediální laboratoře FIS VŠE. V rámci této spolupráce je realizován centrální rozvojový projekt KNOWtilus, jehož cílem je prezentovat vědecký a industriální odkaz v umění postaveném na site specific alternativním divadle. Multimédia jsou nástrojem záznamu, ale především také umělecké tvorby samotné, kdy jsou multimediální prvky insta-

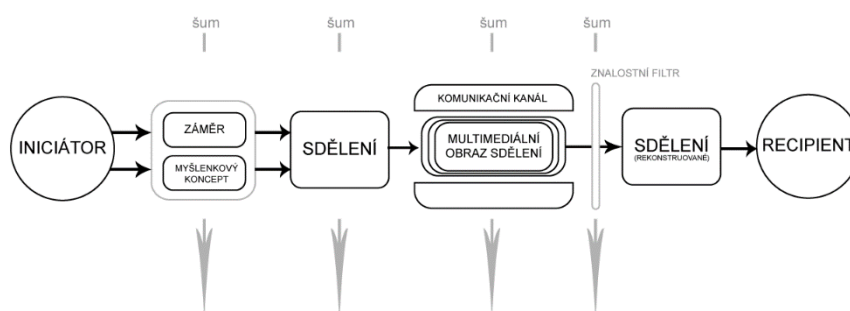
lovány do uměleckých výstupů tak, aby vznikala díla intermediálního charakteru (nestandardní propojení komunikačních forem, které vytváří celek, jehož jednotlivé nástroje jsou na sobě závislé a absence jednoho z nich by narušila sdělovací hodnotu celku) [Sczepanik, 2002]. Projektu se účastní studenti, kteří absolvovali nebo ve stejném čase navštěvují kurz Řízení multimediálních projektů a vzniká tak prostor pro diskusi nad přenášením komunikačních principů z reklamní praxe do alternativního umění. V roce 2014 vyjde monografie popisující tento projekt, ve které publikují kapitolu zaměřenou na analytické zkoumání procesu tvorby tohoto typu alternativního divadla skrze aktuální poznatky z práce na tvorbě modelu multimediální komunikace.

Rozvoj poznatků

V rámci diplomové práce jsem se věnoval rozboru prvků a charakteristik procesu multimediální komunikace skrze hledání transformací a jejich parametrů dle metodiky optimalizace informačních systémů (podnikových procesů) [Adamec, 1993]. Na základě předpokladu, že model bude strukturou odpovídat Shannon-Weaverově modelu a dalším o něho odvozeným, jsem aplikoval do tohoto modelu principy užívání multimediálních nástrojů (obrázek 8). Postupným rozbořením jednotlivých transformací vznikaly závěry v podobě parametrů jednotlivých prvků mající vliv na zvýšení efektu procesu komunikace.

Byly formulovány procesní věty a první pracovní verze modelu multimediální komunikace:

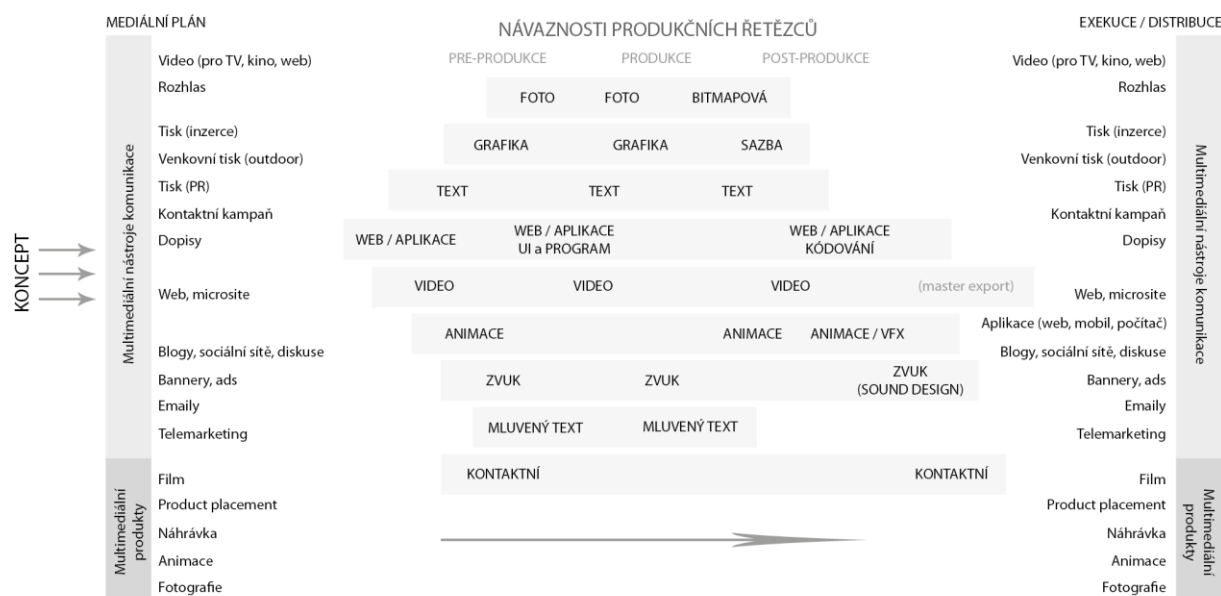
- Iniciátor formuje prostřednictvím záměru a myšlenkového konceptu sdělení.
- Sdělení je kódováno prostřednictvím jednotlivých médií do multimédia, kde je vytvořen multimediální obraz sdělení.
- Multimediální obraz sdělení je přenášen komunikačním kanálem k recipientovi.
- Recipient dekóduje multimediální obraz sdělení do odvozeného sdělení, které vzniká na základě přijetí obsahu.



Obrázek 24: První pracovní verze modelu multimediální komunikace. Zdroj: Vondra, 2011

Integrace produkčního řetězce

Moje další práce v rámci doktorského studia spočívá více než v tvorbě v aplikaci, ověřování a upřesňování modelu a jeho postupné nové formulaci. Během podrobnějšího studia doručovacího, prezentačního a senzorického přístupu k multimédiím jsem zjistil, že pokud hledáme optimalizaci směrem k efektivitě procesu komunikace, nestačí rozebírat pouze proces komunikace samotné, ale je třeba jít hlouběji a zkoumat také procesy, které jeho části přímo tvoří. Jedním z nich je i výrobní řetězec, který v modelu (obrázek 8) odpovídá dějům od prvotní myšlenky iniciátora až po znalostní filtr. Oproti modelu strukturou představuje spíše postup procesního řešení. Transformace a informační potřeby v něm obsažené vstupují zpět do modelu a ovlivňují jeho nastavení. Výrobní (produkční) řetězec bude nakonec tím místem, které bude rozhraním modelu pro optimalizaci informačními a komunikačními systémy, neboť bude mít přímý praktický dopad na vznik komunikačního obsahu a zároveň bude integrovat ekonomické ukazatele. Produkční řetězec platný obecně pro multimediální disciplíny byl studován v rámci grantu OPPA v klíčové aktivitě Komunikace a tvůrčí procesy, jejímž výstupem je pracovní materiál fungující jako e-learningový modul určený k výuce. Základem pro jeho vznik bylo vytvoření přehledu mediatypů (kompletních multimediálních nástrojů) a jejich návaznost na jednotlivé multimediální oblasti (obrázek 9).



Obrázek 9: Přehled médiatypů v reklamě

Zdroj: autor

Médiatypy vyznačené na obrázku 9 jsou odvozené od tvorby reklamy. V rámci projektu KNOWtilus došlo k pojmenování mediátupů použitých v divadelní tvorbě. Ukazuje se, že pro práci s médii je nutno rozlišovat pojem multimédium a pojem mediatyp, který fakticky říká, jakou podobu a zároveň i formu bude multimediální prezentace mít. V návaznosti na Schrammovu teorii komunikace tyto dvě složky nelze oddělovat [Schramm, 1954].

Zjednodušený popis produkčního řetězce:

- **Development** (design projektu): Prvotní myšlenka a zadání, průzkum (mapování prostředí, předběžné informace o zakázce, situační analýza), vývoj (kreativní návrh řešení, návrhy kreativních konceptů), plánování (produkční návrh řešení, návrhy procesu realizace, určení činností, stanovení časových, nákladových a kvalitativních parametrů projektů, mediální plánování, plánování lidských kapacit), vyhotovení a schválení závazného zadání.
- **Pre-produkce**: Dokončení a potvrzení průzkumu (analytické procesy), realizace vývoje (příprava kreativního obsahu), zpracování podrobného plánu (produkční koncepty)
- **Produkce** (realizace projektu): Tvorba obrazového, zvukového a aplikačního materiálu, řešení požadavků na změny a překážek v průběhu realizace, dokumentace průběhu (produkční zprávy), poznámky pro zpracování (skript).
- **Post-produkce**: Shromáždění, selekce, zpracování obrazového a zvukového materiálu, integrace materiálů, kontrola a dokumentace průběhu (produkční zprávy), finalizace do výstupu k předání
- **Exekuce / distribuce**: Použití vytvořeného díla (nasazení) pro komunikaci k cílové skupině, monitorování a hodnocení

Dle první nástrojové roviny zkoumání procesu multimediální komunikace byly přeneseny všechny fáze tvorby multimediálního výstupu za sebe a ukázalo se, že v praxi některé oblasti nevyužívají určité mezi kroky z důvodů, že oblast není tak náročná, aby je bylo třeba provádět standardizovanou formou. Jejich průběh je ponechán intuici a pracovním zvyklostem. Tento fakt však na druhou stranu vede k vyšší míře volnosti a náchylnosti na některé, která stupňuje nahodilost a způsob řešení neformalizovaného úkolu.

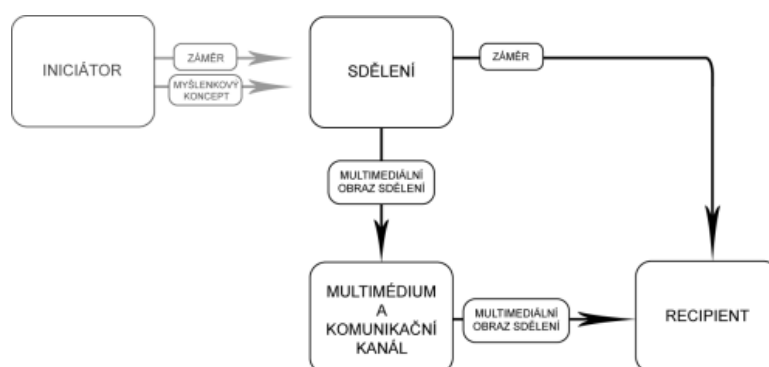
Integrace duálního kódování

Podkladem úpravy na úrovni celého modelu je dále fakt, že prezentovaný model vznikl bez pochopení smyslu duálního kódování a je tedy třeba jej přetvořit s integrací tohoto principu. Model tak bude sice složitější, ale bude komplexněji pokrývat problém, který řeší. Princip duálního kódování, tak jak je

popsán na obrázku 7 bude prostupovat recipientem, multimediálním obrazem a kanály komunikace prezentovanými na obrázku 8. Pro relevantní propojení obou modelů je třeba provést ještě další ověřovací práci. Na základě integrace duálního kódování dojde u každého z užívaných médií ke vzniku dvou parametrů a to „obraz“ a „narace“. Na tomto základě bude možné již během skládání mixu médií u komplexnějších komunikací plánovat dříve budoucí kognitivní efekt a vyvarovat se například přetížení některé z kognitivních rovin příjemce, což by efekt komunikace značně narušilo [Mayer, 2009].

Integrace interaktivity

Z pohledu úprav struktury prvků se ukázalo, tak základním posunem bude rozdělení pozice iniciátora na iniciátora a zpracovatele. Pozorováním praxe bylo zjištěno, že komunikaci věcně v mnoha případech navrhuje a realizuje jiný subjekt, než ten, který komunikaci zadává a má s ní úmysl. Situace, kdy je iniciátor a zpracovatel tentýž subjekt jsou nyní spíše minoritní nebo na úrovni nezávislé tvorby. Tento poznatek navazuje na princip, kdy již v předchozí práci bylo poukázáno na existenci multimediální komunikace v komunikačním prostoru z velké míry nezávisle na iniciátorovi, čímž vzniká požadavek na úplnost a propracovanost produktu (obrázek 10).



Obrázek 10: Oddělené sdělení od iniciátora

Zdroj: autor

S vědomím aktuálních poznatků lze tvrdit, že sdělení je odděleno nejen od iniciátora, ale do velké míry i od zpracovatele. Pokud budeme uvažovat možnosti interaktivity multimediální komunikace, narazíme na problém, že každý multimediální produkt vzniká dle svého produkčního řetězce v komunikačním vákuu a do prostoru komunikací je v okamžiku dokončení a čase požadovaného nasazení vypuštěn. Díky tomuto předpokladu je jediná možnost zpětné vazby na úrovni interaktivní komunikace dle Rothwella zmíněné výše, kdy recipienti s časovou prodlevou vyjádří verbálně své pochopení a s další časovou prodlevou zpracovatel vytváří novou multimediální odpověď. Časové prodlevy mají dopad na formování efektu aktuálně nasazené komunikace a komunikace, která ji bude v budoucnu nahrazovat, upřesňovat nebo rozvíjet. V tomto kontextu zatím multimédia nenabízejí takové technologické řešení, které by bylo schopné realizovat efektivní transakční komunikaci.

Díky projektu KNOWtilus a prvotní práci s intermédii jsou do modelu vneseny nové otázky spojené s interaktivitou a to: Jaký má význam interaktivita recipienta s komunikací? Jaký význam má interaktivita médií mezi sebou? Jak je možné efektivně regulovat interaktivitu tvůrce s komunikací?

Závěr

Model multimediální komunikace má šanci v návaznosti na studium modelů komunikace a přenosu informací skrze multimédia poskytnout funkční rámec pro tvorbu multimediálních nástrojů. Pokud bude jeho formulace dostatečně podložena a ověřena v různých oblastech, kde jsou multimédia využívána za různým účelem, vznikne metodická konstrukce, která usnadní a urychlí, odborníkům i laikům přípravu tvorby multimediálního obsahu tak, aby výstup efektivněji naplňoval požadovaný záměr. V kontextu digitalizace a rozšíření užívání multimediálních nástrojů by vznik modelu umožnil vytvořit konkrétní metodiku nebo doporučení pro tvorbu obsahu určeného k osobním, komerčním nebo jiným sdělovacím účelům. Neskromným cílem je i obecné zvýšení mediální gramotnosti.

Částečně opomíjenou velkou oblastí a otázkou, před kterou práce stojí, je interaktivita uvažovaná nejen jako nástroj zpětné vazby, ale také jako nástroj podporující efekt komunikace skrze zapojení recipienta.

Aktuální model je postaven na lineárním principu, což není do budoucna ve vztahu k pokroku technologií a odklonu od masových komunikací k mířeným dostačující. Interakce tedy bude další oblastí, která bude studována a do modelu integrována.

Rozsah práce vybízí k sestavení pracovního týmu, který by se spolupodílel na ověřování a formulaci modelu. Každý z jeho členů by se zabýval konkrétní oblastí a propojoval by ji s obecným rámcem. Postupově stojí práce před velkým úkolem a to je integrovat pracovní model s produkčním řetězcem, modelem duálního kódování a interakcemi recipienta, nástrojů a tvůrce. Před provedením těchto úkonů bude nutné provést rešerši různých vědeckých prací zaměřených na spojování různých přístupů a formulaci účelových modelů. Tyto práce jsou často postaveny na protnutí teoretického modelu a konkrétní oblasti z praxe. Poslouží jako metodický rámec zkušeností se slučováním konceptuálních poznatků.

Literatura

ADAMEC, Stanislav, TRHOŇ, Daniel, 1993. *Člověk, data, informace*. 1. vyd. Praha: Vysoká škola ekonomická v Praze, 1993. ISBN 80-707-9891-2.

BARNLUND, D. C., 2008. *A transactional model of communication*. In: C. D. Mortensen (Eds.), *Communication theory* (2nd ed., pp47-57). New Brunswick, New Jersey: Transaction, 2008.

CORMAN, Steven R., Angela TRETHERWEY a Bud GOODALL. 2007. *A 21st Century Model for Communication in the Global War of Ideas: From Simplistic Influence to Pragmatic Complexity*. 2007, s. 18. Dostupné z: <http://csc.asu.edu/wp-content/uploads/pdf/114.pdf>

CRAIG, Robert T., 1999. *Communication theory as a field*. *Communication Theory*. 1999, 9, 2, s. 199-161.

HARDMAN, Lynda, 2005. *Canonical Processes of Media Production*. In: GENERAL CHAIRS, Frank Nack a SIGGRAPH and SIGMULTIMEDIA] [SPONSORS.Proceedings of the ACM Workshop on Multimedia for Human Communication Hilton (Singapour), November 11, 2005. New York, N. Y: ACM Press, 2005. ISBN 159593247x.

HIGGINS, Dick, 1966. *Intermedia*, Something Else Newsletter 1, 1966

HONTHANER, Eve Light, 2010. *The complete film production handbook*. 4th ed. Burlington, MA: Focal Press, xviii, 524 p. ISBN 02-408-1150-X.

CHECKLAND, Peter, 1999. *Soft systems methodology: a 30-year retrospective*. New York: John Wiley, c1999, 1 v. (various pagings). ISBN 04-719-8606-2.

LASSWELL, HAROLD, 1948, "The Structure and function of communication in society" in *The Process and Effects of Mass Communication*. Wilburn Schramm and Donald F. Roberts, eds., Urbana: University of Illinois Press, 1971, 84-99

MAYER, Richard E. a Roxana MORENO, 2003. *Nine Ways to Reduce Cognitive Load in Multimedia Learning*. *EDUCATIONAL PSYCHOLOGIST*. 2003, 38(1), č. 1, 43–52.

MAYER, Richard E., 2009. *Multimedia learning*. 2nd ed. Cambridge: Cambridge University Press, 2009. ISBN 05-215-1412-6.

MCLUHAN, Marshall, 2011. *Jak rozumět médiím: extenze člověka*. 2., rev. vyd. Překlad Miloš Calda. Praha: Mladá fronta, 2011, 399 s. ISBN: 978-80-204-2409-9

PAIVIO, Allan, 1986. *Mental representations: a dual coding approach*. Oxford. England: Oxford University Press, 1986. ISBN-10: 0195066669

ROTHWELL, J. Dan, 2009. *In the Company of Others : An Introduction to Communication*. 3. New York: Oxford University Press, 2009. 544 s. ISBN 9780195336306.

SHANNON, C. E., 1948. *A Mathematical Theory of Communication*. The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, July, October, 1948.

SCHRAMM, W., 1954. *How communication works*. In W. Schramm (Ed.), *The process and effects of communication* (pp. 3-26). Urbana, Illinois: University of Illinois Press, 1954.

SCZEPANIK, Petr, 2002. *Intermedialita*. Cinepur [online]. roč. 2002, č. 22 [cit. 2014-01-25].

VAUGHAN, Tay, 2008. *Multimedia: making it work*. 7th ed. New York: McGraw-Hill, c2008, xiii, 514 p. ISBN 00-722-6451-9.

VONDRA, Z., 2011. *Multimediální komunikace*. Praha, 2011. Diplomová práce. Vysoká škola ekonomická v Praze. Vedoucí práce Stanislav Horný.

communicationtheory.org, 2010 *Berlo's model*. [online]. [cit. 2014-01-20]. Dostupné z: <http://communicationtheory.org/berlos-smcr-model-of-communication/>

communicationtheory.org, 2010 *Lasswell's model*. [online]. [cit. 2014-01-20]. Dostupné z: <http://communicationtheory.org/lasswells-model/>

JEL Classification: D83.

Summary

Multimedia communication model preparation

The article presents the current state of preparations of multimedia communication model launched in 2011 in connection with the establishment of Graphic and Multimedia Laboratory of the Faculty of Informatics and Statistics of University of Economics, Prague. The requirement for the creation of a model was based on a lack of conceptual methodological materials for work with multimedia that would describe communication as a whole process without direct relation to one particular tool. There exist such general concepts, but they each describe specific multidisciplinary phenomena or problem, but not the communication as a whole process.

As a result of digitization, there are available in almost all multimedia tools to the general public and thus becomes a key tool for their usage strong conceptual level that is applicable in multiple areas of media at once. In practice, this means that for example advertising agencies want to increase the value of the idea concepts of their own production instead of producing media tools that are being often outsourced.

Preparations for multimedia communication model are based on traditional models and theories. The work considered the fact, that communication is an interdisciplinary area. The article summarizes the current background, methods and actual process of model building. There are presented two levels of authentication: a common tool level and a society level that is derived from a variety of areas using multimedia tools. In the third part are presented new findings that push the model further in the development. There are also presented new tasks to work on the model.

Key words: multimedia, multimedia communication, multimedia design, communication model, dual coding, interactivity, production chain.

Vývoj nástroje pro diskrétní simulaci

Karel Charvát

charvat.k@seznam.cz

Doktorand oboru ekonometrie a operační výzkum

Školitel: prof. Ing. Josef Jablonský, CSc. (jablon@vse.cz)

Abstrakt: Článek je věnován programu pro diskrétní simulaci nazvanému SIMULANT. Jedná se o volně dostupný doplněk pro aplikaci MS Excel vytvořený ve Visual Basic for Applications (VBA). První verzi byla vyvinuta v rámci mé diplomové práce. Vývoj probíhá i nadále. Důvodem pro další vývoj doplňku je v současnosti především nízká vybavenost českých firem komerčními simulačními programy a nedostatečná nabídka uživatelsky přívětivého freewaru nebo open source softwaru pro diskrétní simulaci

Představení doplňku SIMULANT v tomto článku obsahuje základní informace o jeho možnostech, uživatelském rozhraní, dostupných prvcích, parametrech jednotlivých prvků a obsluze tohoto doplňku. Součástí popisu doplňku je i ukázka záznamu simulačních událostí, který slouží k automatickému generování výsledkových zpráv.

Další vývoj doplňku SIMULANT zahrnuje probíhající i plánované přidání nových možností týkajících se přiřazování zdrojů, proměnlivosti určitých parametrů modelu v čase i řízení průchodu entit modelem. Cílem je také celkové zvýšení výkonu programu a vydání doplňku jako open source.

Klíčová slova: diskrétní simulace, SIMULANT, software, open source, Excel.

Úvod

Simulační modelování vzniklo z metody Monte Carlo. Hlavní myšlenkou metody Monte Carlo je řešení pravděpodobnostních i deterministických úloh pomocí statistického experimentu. Někteří autoři považují pojmy simulace a Monte Carlo za shodné, jiní je rozlišují podle využití. Dle jejich členění je metoda Monte Carlo využívána k řešení úloh, ve kterých nehraje roli dynamické chování, zatímco simulace ke studiu dynamických systémů, tedy takových, ve kterých je podstatným faktorem čas (Dlouhý et al., 2007). Simulaci můžeme dělit na několik typů dle způsobu zachycení času a stavů systému. Za diskrétní považujeme takovou simulaci, kdy změny systému nejsou zaznamenávány průběžně, ale pouze při výskytu určitých událostí, které mohou nastat v libovolném okamžiku ve sledovaném časovém úseku, a jsou z hlediska zkoumání daného systému relevantní. Časové úseky, ve kterých se žádná ze sledovaných událostí nevyskytuje, jsou v diskrétní simulaci přeskočeny. Interval mezi výskytem těchto událostí obvykle nebývá stejně dlouhý a jeho délka často má náhodný charakter.

Simulace byla v počátku využívána převážně v technických oborech či pro modelování chemických procesů. Později se však rozšířila i do ekonomické oblasti, neboť možnost předvídat důsledky nejruznějších vnitřních i vnějších změn na chování určitého systému je velice užitečná i při analýze ekonomických procesů. (Dlouhý et al. 2007). Možnost simulovat změny systému pomocí modelu bez nutnosti zasahovat do reálného systému a porovnávat nejruznější varianty činí simulaci vhodným podpůrným nástrojem pro manažerské rozhodování. S diskrétní simulací se nejčastěji setkáváme právě v ekonomické oblasti.

V následujících kapitolách bude čtenářům představen můj nástroj pro diskrétní simulaci, jehož první verzi jsem vytvořil v rámci diplomové práce. Vývoj tohoto nástroje probíhá i nadále. Jedná se o doplněk SIMULANT pro MS Excel. První část je věnována důvodům, které vedly ke vzniku a dalšímu vývoji tohoto nástroje a následující kapitola je zaměřena na popis doplňku SIMULANT. Cílem poslední části je seznámení čtenáře s dalším probíhajícím i plánovaným vývojem doplňku SIMULANT.

Důvody pro vznik a vývoj doplňku SIMULANT

S operačním výzkumem, mezi jehož disciplíny patří i simulační modelování, se v dnešní době běžně setkávají studenti na matematicky, technicky i ekonomicky zaměřených studijních programech vysokých škol. Dle zaměření studia bývá kladen různý důraz na teoretickou nebo aplikační stránku. Využití

získaných znalostí a dovedností při řešení reálných problémů se ve většině případů neobejde bez vhodného softwaru. Během studia se studenti seznamují s nejrůznějšími druhy modelovacího, optimalizačního i simulačního softwaru. Produktů zaměřených na matematické programování, či produktů pro simulaci je v dnešní době na trhu velké množství. Obsáhlý přehled současných komerčních simulačních nástrojů můžete nalézt například v (Swain, 2013).

Nedostupnost vhodného softwaru ve firmách však bývá pro absolventy v některých případech hlavní překážkou, která omezuje jejich možnosti uplatnění metod operačního výzkumu v praxi. Ačkoliv se studenti ve výuce setkávají i s nástroji, které patří mezi nejkvalitnější komerční produkty, a tudíž by bylo možné očekávat jejich hojné využití v praxi, většina českých firem takovými nástroji nedisponuje. Nepodařilo se mi nalézt žádnou konkrétní statistiku týkající se využití tohoto softwaru, ale můj úsudek je založen na skutečnosti, že požadavky znalosti produktů pro matematické programování jako například MPL, GAMS, AMPL, AIMMS, LINGO, CPLEX nebo GUROBI, či produktů pro diskrétní simulaci jako SIMUL8 a SIMPROCESS se na českých portálech s nabídkami zaměstnání vyskytují velice vzácně.

Za hlavní důvod relativně malého rozšíření těchto nástrojů ve firmách nepovažuji nedostatek problémů, pro jejichž řešení jsou dané nástroje vhodné ani nedostatek pracovníků, kteří takové nástroje ovládají. Příčinu vidím především ve vysoké pořizovací ceně. Jako příklad uvedu cenu jedné licence produktu SIMUL8, která činí 995 amerických dolarů pro verzi Basic a 4995 dolarů pro verzi Professional (SIMUL8 Corporation, 2014). Taková cena je nejspíše přijatelná pro firmy opakovaně nabízející služby spojené s různými druhy optimalizace či simulace, ale méně přijatelné pro firmy zvažující jednorázová řešení určitého problému vlastními silami.

Východiskem z této situace může být využití nástrojů z kategorie freewaru či open source. V posledních letech vzrůstá počet i aktivita iniciativ, které freeware či open source pro operační výzkum vyvíjejí. Především oblast matematického programování je dnes pokryta širokou nabídkou produktů. V minulosti freeware či open source nástroje pro matematické programování vznikaly především v podobě knihoven pro různé programovací jazyky. Dnes už však v této kategorii můžeme narazit i na produkty s grafickým uživatelským rozhraním (COIN-OR, 2012). Nejvýznamnější organizací v oblasti open source softwaru pro operační výzkum je výše zmiňovaná COIN-OR, jejíž nástroje jsou dále využívány i mnohými autory softwaru působícími mimo tuto iniciativu.

Freeware či open source nástroje pro diskrétní simulaci se však zatím vyskytují téměř výhradně v podobě různých balíčků či knihoven pro programovací jazyky. Přehled mnoha příkladů takových nástrojů můžete nalézt v (List of discrete event simulation software, 2014). Z open source nástrojů uvedených v tomto seznamu je zřejmě nejznámějším balíček SimPy pro jazyk Python. Využití takových knihoven či balíčků nabízí vysokou flexibilitu, ale klade relativně velké nároky čas a programátorské schopnosti uživatele. Pro studenty, kteří se během výuky seznámí například s nástroji SIMUL8 či SIMPROCESS, tedy programy vyučovanými v rámci předmětu simulační modely na VŠE, může být v praxi velice problematické nalezení volně dostupného nástroje umožňujícího pracovat se simulačními modely způsobem, který si osvojili během studia. K překonání tohoto problému bych rád přispěl pomocí dalšího rozvoje svého vlastního nástroje pro diskrétní simulaci, jehož první verze vznikla v rámci má diplomové práce. Jedná se o doplněk SIMULANT pro MS Excel, pro jehož vývoj jsem použil Visual Basic for Applications (VBA).

Skutečnost, že je MS Excel komerčním produktem, neopovažuji při vývoji volně dostupného doplňku za podstatnou nevýhodu. Alespoň základní znalost MS Excel patří na trhu práce k nejčastěji požadovaným počítačovým dovednostem napříč odvětvími, obzvláště potom na analytických pozicích. Vybavenost firem tímto produktem je tedy velice vysoká. Doplněk SIMULANT pro MS Excel je tedy nástrojem dostupným široké skupině uživatelů. Původně byl vyvíjen především jako snadno ovladatelný nástroj sloužící především pro výukové účely, umožňující ilustraci základních principů diskrétní simulace. Cílem dalšího vývoje doplňku je především přidání nových funkcí, které zvýší využitelnost při řešení praktických úloh a dále potom proměna nástroje v open source.

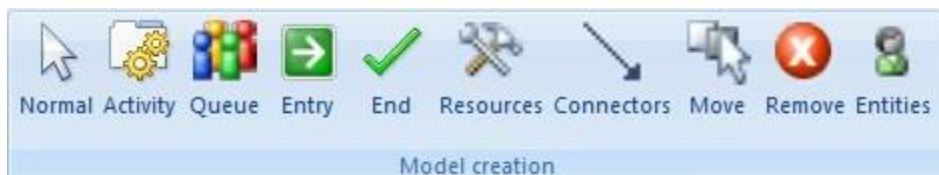
Seznámení s doplňkem SIMULANT

SIMULANT je nástroj pro diskretní simulaci, který umožňuje vytvářet simulační modely v grafickém režimu. Jedná se o doplněk pro MS Excel vytvořený v jazyku VBA (Visual Basic for Applications). První verze doplňku SIMULANT vznikla v rámci diplomové práce, kterou jsem v roce 2009 obhájil na Vysoké škole ekonomické v Praze, katedře ekonometrie. Doplňek SIMULANT ke své činnosti vyžaduje MS Excel 2007, nebo novější. Ve starších verzích MS Excel není jeho fungování možné, a to především kvůli zvolenému uživatelskému rozhraní, kterým je pás karet poprvé se vyskytující ve verzi 2007.

Doplňek nabízím ke stažení a instalaci zdarma a účel jeho využívání neomezují žádnými pravidly. Uživatel může doplněk SIMULANT využívat pro jakékoliv simulace, pro které jsou funkce a výkon doplňku dostačující. Jedná se tedy o nástroj, který v současnosti patří do kategorie freeware. První verze doplňku obsahovala uživatelské rozhraní v českém jazyce, verze 1.0.3 existovala ve dvou jazykových variantách, české a anglické. Pozdější verze jsou vyvíjeny výhradně s anglickým uživatelským rozhraním.

Nejnovější verzi doplňku můžete stáhnout ze stránek <http://simulant-addin-en.webnode.cz/>. Instalaci je možné provést standardním způsobem instalace doplňků pro aplikaci Excel, případně pomocí dialogu vyvolaného otevřením souboru s doplňkem.

Hlavní ovládací prvky doplňku se nacházejí na kartě SIMULANT, která na pásu karet po instalaci doplňku přibude. Dostupnost jednotlivých ovládacích prvků je ovlivněna tím, v jakém režimu se model v danou chvíli nachází. Na kartě SIMULANT je možné zvolit typ vkládaného prvku a následně umisťovat prvky do modelu kliknutím na příslušné místo podkladu. Stejně snadno lze definovat i vazby mezi jednotlivými prvky modelu.



Obrázek 1: Tlačítka pro tvorbu modelu

V doplňku SIMULANT jsou k dispozici prvky entita, vstup, fronta, aktivita, výstup, zdroj a spojnice. Názvosloví jednotlivých prvků modelu se mezi jednotlivými simulačními programy samozřejmě může lišit. Většina z těchto typů prvků má určité nastavitelné atributy.

Entita (Entity) je typ prvku, který prochází simulačním modelem. V Doplňku SIMULANT je možné definovat více typů entit.

Vstupy (Entry) slouží jako objekty, kterými entity přicházejí do modelu. Pro vstup je nejdůležitějším nastavitelnou vlastností interval mezi příchody jednotlivých entit. K dispozici je několik náhodných rozdělení, kterými se interval mezi příchody může řídit. Jedná se o exponenciální, rovnoměrné, celočíselné rovnoměrné, trojúhelníkové a normální rozdělení. Pro každé z těchto rozdělení je samozřejmě možné nastavit jeho parametry. Ke generování náhodných čísel je využita funkce *Rnd()* jazyku VBA. Následuje transformace vygenerovaných náhodných čísel na hodnoty z příslušných náhodných rozdělení. Principy těchto transformací jsou popsány např. v (Dlouhý et al., 2007) nebo (Banks, 1998).

Fronty (Queue) slouží jako objekty, ve kterých entity mohou vyčkávat na uvolnění místa v objektu, který za frontou následuje. Fronta se může řídit jedním z režimů FIFO (first in first out), LIFO (last in first out), nebo SIRO (select in random order). Pro frontu je dále možno nastavit její maximální možnou délku. Také je možno určit, že pořadí entit opouštějících frontu se bude řídit prioritami. V takovém případě se priorita stává primárním kritériem a výše uvedené možnosti slouží jako pomocné kritérium v případě rovnosti priorit. Priority je možno entitám přiřadit při vstupu do modelu.

Aktivity (Activity) reprezentují činnosti, kterými entity během simulace procházejí. Nejdůležitější nastavitelnou vlastností je náhodná doba trvání aktivity. Tato doba se může řídit některým z náhodných rozdělení, která byla uvedena u objektu *vstup*. Další důležitou vlastností je násobnost dané aktivity. Je-li

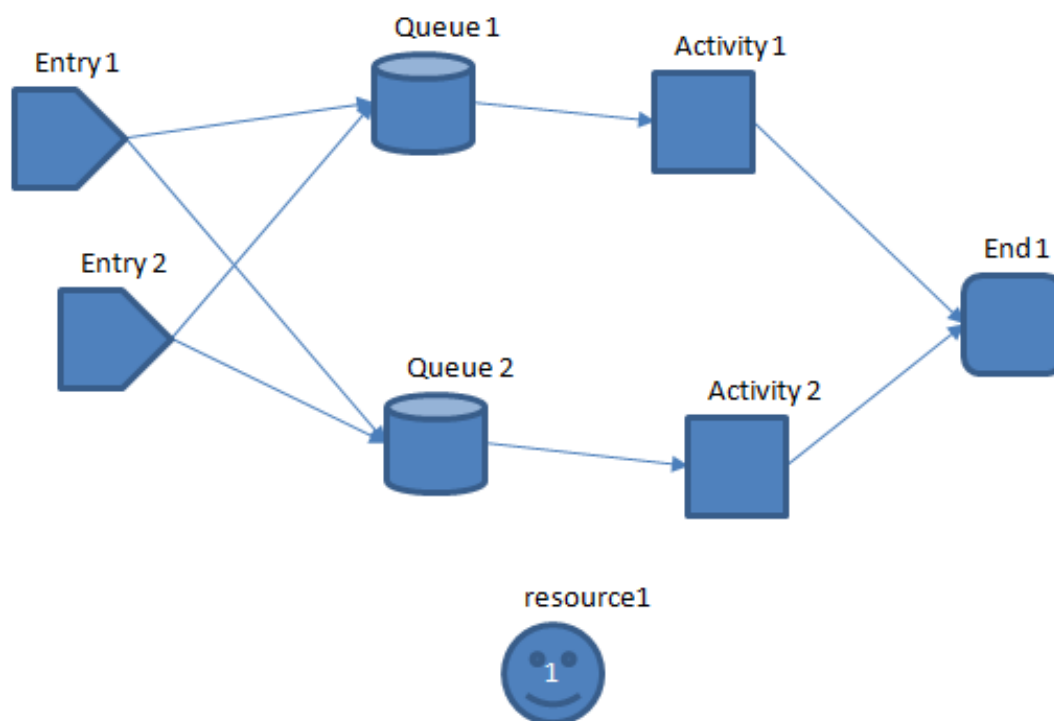
násobnost větší než 1, graficky je aktivita reprezentována stále jen jedním objektem, ale do simulačního modelu je zařazena vícekrát paralelně, v nastaveném počtu.

Zdroje (Resource) jsou prvky modelu, které mohou být v různém počtu vyžadovány k vykonávání určitých aktivit. Pro vykonávání aktivity je možné požadovat jeden, či více typů zdrojů v libovolném počtu. Některé typy zdrojů mohou být vyžadovány více různými aktivitami. Aktivita vyžadující zdroje nemůže být zahájena, dokud jí nejsou všechny požadované zdroje v potřebném počtu přiřazeny. Po dobu vykonávání aktivity jsou zdroje přiřazené této aktivitě blokovány. Dokud není daná aktivita ukončena a zdroje uvolněny, nemohou být využívány v rámci jiné aktivity.

Výstup (End) je objekt, kterým entity opouštějí model. Tento typ objektu nemá žádné nastavitelné parametry ovlivňující chod simulace.

Spojnice (Connectors) slouží pro definování vazeb mezi jednotlivými objekty. Vstupům, frontám a aktivitám je možno pomocí spojnic přiřazovat následníky a vytvářet tak strukturu modelu. Jednomu objektu může být přiřazeno více následníků.

Následující obrázek zachycuje grafickou podobu jednoduchého ukázkového modelu vytvořeného pomocí doplňku SIMULANT.



Obrázek 2: Ukázka grafického znázornění modelu

Nastavování vlastností jednotlivých prvků modelu je možné pomocí formulářů, které lze zobrazit kliknutím na příslušný prvek. Formulář otevřený kliknutím například na objekt pojmenovaný *Activity 1* je zachycen na **obrázku 3**. Každý typ prvku modelu má určité výchozí nastavení parametrů, které je použito, pokud uživatel hodnoty nezmění. Pro intervaly mezi příchody i doby trvání aktivit je výchozím nastavením exponenciální rozdělení se střední hodnotou 10 minut.

Pro vstupy, fronty a aktivity je možno vybrat některé z pravidel, dle kterého entity opouštějící daný objekt volí, ke kterému z jeho následníků se přesunou. Určení pravidel pro volbu následníků je možné provést formulářem, který lze otevřít pomocí tlačítka *Routing* v nastavení daného objektu. Výchozím pravidlem je přiřazení stejné pravděpodobnosti každému z následníků určitého objektu. Každému vstupu, frontě nebo aktivitě můžeme určit jedno z následujících pravidel.

Obrázek 3: Ukázka ovládacího formuláře

Pravděpodobnost: každému z následníků určitého objektu je možno přiřadit pravděpodobnost, se kterou je volen.

Dle typu: pomocí tohoto pravidla můžeme určit, kterým následníkům budou jednotlivé typy entit směřovat.

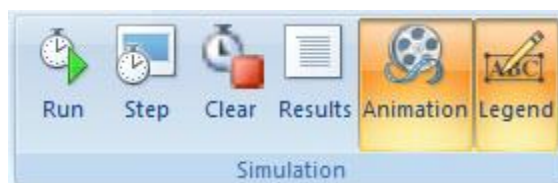
Nejkratší fronta: entita volí následníka s nejkratší délkou fronty. Výstupy a volné replikace aktivit jsou považovány za následníky s nulovou délkou fronty.

Preference: entita opouštějící objekt vybírá následníka dle seznamu preferencí. Jako následník je vybrána položka nejvýše v seznamu, která má volnou kapacitu.

ID	Name	Probability
f_1	Queue 1	0,5
f_2	Queue 2	0,5

Obrázek 4: Nastavení pravidel pro volbu následníka

Pro samotný běh je možné nastavit simulační čas, ve kterém bude simulace ukončena. Simulaci je také možno kdykoliv v průběhu pozastavit a následně opět spustit. Další možností využitelnou při chodu simulace je krokování. Při krokování dochází k posunu simulačního času na okamžik nejbližší události, zpracování událostí, které se v tomto čase odehrají a následnému automatickému pozastavení.



Obrázek 3: Tlačítka pro řízení chodu simulace

Kdykoliv během pozastavení simulace nebo po jejím ukončení je možno vygenerovat výsledkovou zprávu. Z jednoho běhu simulace je možno vygenerovat více výsledkových zpráv v různých simulačních časech. Výsledková zpráva obsahuje charakteristiky jednotlivých prvků modelu získané z aktuálního běhu simulace. V případě vstupů se jedná například o počet entit vygenerovaných vstupem. Výsledky týkající se front obsahují například údaje o průměrných dobách čekání v dané frontě a průměrné délce fronty. Výsledky týkající se aktivit informují o počtech obslužených entit a využití dané aktivity. U zdrojů je hlavní informací jejich vytíženost. U jednotlivých výstupů můžeme zjistit například průměrné doby setrvání entit v systému.

V průběhu simulace je možno informace o chování modelu získávat také pomocí jednoduché animace znázorňující pohyby entit a pomocí průběžně aktualizovaných čísel zobrazovaných na jednotlivých objektech. V případě vstupů tato čísla udávají počet entit, který byl do aktuálního času daným vstupem vygenerován, v případě front a aktivit počet entit, které se v daném objektu momentálně nalézají, u výstupů celkový počet entit, které daným výstupem opustily model. U zdrojů je tímto způsobem možno sledovat aktuálně volný počet jednotek daného zdroje.

Doplňek využívá ke své činnosti několik různých listů. Tyto listy jsou při tvorbě nového modelu zkopírovány z doplňku do aktivního sešitu aplikace Excel. Jeden list slouží jako součást uživatelského rozhraní a obsahuje podklad pro grafickou podobu modelu. Ostatní listy jsou programem využívány k ukládání různých typů dat. Tyto listy jsou skryty, ale proti zobrazení nejsou nijak chráněny. Při pozastavení simulace, nebo po jejím ukončení je tedy možné údaje obsažené v těchto listech prohlížet. Informace obsažené v těchto listech však slouží především samotnému programu a jejich znalost není pro uživatele při provádění simulace nezbytná. Mohou však být užitečné při snaze porozumět použitým simulačním algoritmům. Především pokročilejší uživatelé potom mohou zobrazit například informace v listu *events*, který zaznamenává všechny simulační události, a využít je k výpočtu charakteristik, které překračují rámec standardní výsledkové zprávy.

Základním principem, na kterém je simulační algoritmus použitý v doplňku založen, je *časový krok proměnlivé délky*. Myšlenka tohoto principu je velice obecná a existuje mnoho naprosto odlišných způsobů, jak tento princip do simulačních programů implementovat. V literatuře (Banks, 1998) jsou obecně popsány příklady různých přístupů k tomuto principu, avšak konkrétní realizace v simulačním softwaru obvykle bývá pro uživatele černou skříňkou. Pro doplňek SIMULANT jsem navrhl vlastní přístup, jehož podrobný popis můžete nalézt v (Charvát, 2009). Při tvorbě algoritmu jsem nejprve definoval typy událostí, které mohou během simulace nastat a vytvořil pravidla jak tyto události zpracovat, aby simulační modely vykazovaly požadované chování. Algoritmus pracuje s událostmi, které vyvolávají posun simulačního času i událostmi, které vznikají v důsledku jiných událostí a stavu modelu v daném okamžiku. Určitá událost tedy může vyvolat i několik jiných událostí, které musí být v daném simulačním čase zpracovány a v mnoha případech se to opravdu stává. Algoritmus je navržen tak, aby fungoval pro model s libovolným počtem prvků i libovolným uspořádáním.

Obrázek 6 slouží jako ukázka záznamu událostí na listu *events*. Události nastaly během jedné velice krátké simulace, jejíž čas byl nastaven na 20 minut. Pro simulaci byl použit model se strukturou znázorněnou na **obrázku 2**. Všechny elementy modelu kromě aktivit byly ponechány ve výchozím nastavení. Pro obě aktivity byl nastaven požadavek na 1 jednotku zdroje nazvaného *resource1*.

Time	Location	Replication	Entity	Type of event	Resource	Number
5,5607	z_2		e_1	Creation		
5,5607	f_1		e_1	Arrival		
5,5607	f_1		e_1	Departure		
5,5607	a_1	1	e_1	Arrival		
5,5607	a_1	1	e_1	Awaiting resources		
5,5607	a_1	1		Resource allocation	resource1	1
5,5607	a_1	1		Resource arrival	resource1	1
5,5607	a_1	1	e_1	Start		
10,5566	z_1		e_2	Creation		
10,5566	f_1		e_2	Arrival		
11,3407	z_2		e_3	Creation		
11,3407	f_2		e_3	Arrival		
11,3407	f_2		e_3	Departure		
11,3407	a_2	1	e_3	Arrival		
11,3407	a_2	1	e_3	Awaiting resources		
16,0008	a_1	1	e_1	Activity end		
16,0008	a_1	1		Resource release	resource1	1
16,0008	a_1	1	e_1	Departure		
16,0008	k_1		e_1	Arrival		
16,0008	f_1		e_2	Departure		
16,0008	a_1	1	e_2	Arrival		
16,0008	a_1	1	e_2	Awaiting resources		
16,0008	a_2	1		Resource allocation	resource1	1
16,0008	a_2	1		Resource arrival	resource1	1
16,0008	a_2	1	e_3	Start		
16,2646	a_2	1	e_3	Activity end		
16,2646	a_2	1		Resource release	resource1	1
16,2646	a_2	1	e_3	Departure		
16,2646	k_1		e_3	Arrival		
16,2646	a_1	1		Resource allocation	resource1	1
16,2646	a_1	1		Resource arrival	resource1	1
16,2646	a_1	1	e_2	Start		

Obrázek 6: Ukázka zaznamenaných událostí

Záznam událostí v této podobě slouží pro automatické generování výsledkové zprávy. Z prvních řádků můžeme vyčíst například, že v simulačním čase 5,56 minut byla ve vstupním bodě, kterému je přiděleno id *z_2* (*Entry 2*), vygenerována entita s id *e_1*. Tato entita ve stejném okamžiku vstoupila do fronty *f_1*, ze které se ihned přesunula do aktivity *a_1*. Pro aktivitu *a_1* byl vyžádán zdroj *resource1*, který byl aktivitě ihned přidělen. Aktivita *a_1* byla zahájena. Následuje zjištění nejbližší další události a posun hodin do simulačního času 10,56 minut. Následující řádky bychom mohli interpretovat obdobným způsobem.

Výsledky například pro frontu nazvanou *Queue 1*, které bylo přiděleno id *f_1*, automaticky generované na základě záznamů zobrazených na **obrázku 6** mají následující podobu.

Queue name:	Queue 1	By entity type
		basic
Queue length		
Actual:	0	0
Average:	0,33	0,33
Maximal:	1	1
Total entred	2	2
Non zero queuing times	1	1
Queuing time		
Average:	2,72	2,72
Maximal:	5,44	5,44
Standard deviation:	2,72	2,72

Obrázek 7: Ukázka výsledků pro frontu

Další rozvoj doplňku

Doplňek SIMULANT byl vytvářen převážně jako snadno ovladatelný výukový nástroj. V současnosti probíhají změny, které mají za cíl rozšířit funkčnost doplňku a dále změny směřující k publikaci doplňku jako open source.

Funkční změny

I přes primární výukové zaměření původní verze umožňovala simulaci relativně jednoduchých reálných systémů. Na základě seznámení s několika případovými studiemi publikovanými například v rámci diplomových prací, které zahrnují simulace prováděné v jiných nástrojích, jsem však identifikoval hlavní nedostatky doplňku SIMULANT, které omezují provádění komplexnějších simulací. Zaměřuji se tedy na rozšíření funkčnosti nástroje především prostřednictvím odstranění nedostatků, které jsou omezující nejvíce. V následujících odstavcích uvedu příklady několika z nich.

V první řadě se jedná o přiřazování zdrojů aktivitám. V původní podobě dochází k přidělení zdroje pouze pro vykonání určité aktivity, po jejímž ukončení je zdroj opět uvolněn. Přidělování zdrojů hodlám rozšířit o možnost přidělit zdroj určité entitě, které tento zdroj zůstane přidělen při průchodu několika různými aktivitami. Takovýto způsob práce se zdroji má široké využití. Jako jednoduchá ukázka poslouží např. zkoumání provozu čerpací stanice. V takovém systému můžeme za jeden ze zdrojů považovat např. čerpací stojany. Definujeme-li v tomto systému jako zkoumané aktivity tankování, nákup doplňkového zboží v obchodě čerpací stanice a placení, je vhodné považovat stojany za blokové daným zákazníkem během všech uvedených aktivit.

Další vyvíjenou změnou je možnost nastavení některých parametrů modelu, například intervalů mezi příchody entit, jako proměnlivých v čase. Tato změna rovněž odráží reálné chování mnoha systémů, jejichž vytiženost se mění v závislosti na denní době, dni v týdnu, nebo jiném faktoru.

Za jednu z budoucích priorit považuji možnost regulovat průchod entit následujícími částmi modelu na základě jejich předchozí cesty modelem. Taková možnost je užitečná ve všech případech, kdy průchody entit určitými aktivitami způsobí změnu vlastností či chování entity a tato změna ovlivní následující cestu.

Mou snahou je rovněž celkové zvýšení výkonu doplňku SIMULANT, který v současnosti znatelně klesá s rostoucím rozsahem modelu. V době tvorby tohoto doplňku v rámci diplomové práce jsem upřednostňoval jednoduchost a názornost před efektivností algoritmů a jejich implementace. Při snaze o zvýšení praktické použitelnosti však nyní na efektivnost kladu větší důraz.

Cílem výše uvedených změn není vznik nástroje, který by byl svými možnostmi schopen konkurovat komerčním produktům, nebo flexibilitou simulačním knihovnám programovacích jazyků. Snažím se o vytvoření nástroje, který absolventům základních kurzů simulací umožní co největší uplatnění simulačních dovedností v praxi.

Vydání doplňku jako open source

Kromě přidání nových funkcí je mým cílem vydání doplňku SIMULANT jako open source. Open source je software s otevřeným zdrojovým kódem. Pojem otevřenost zde zahrnuje technickou i legální dostupnost. Open source software je tedy šířen pod některou z licencí, které uživatelům umožňují při dodržení určitých podmínek zdrojový kód využívat, modifikovat i šířit (Open-source software, 2014). Za open source v užším smyslu je považován software šířený pod některou z licencí vyhovujících požadavkům Open Source Initiative (OSI). Kromě pojmu open source existuje i pojem Free Software (svobodný software), který má s pojmem open source i přes určité odlišnosti mnoho společných vlastností. Klíčovou autoritou v oblasti svobodného softwaru je Free Software Foundation (FSF). Podrobné vysvětlení rozdílů mezi Open Source a Free softwarem překračuje rámec této práce, ale obecně lze konstatovat, že myšlenky, které šíří OSI, jsou orientovány převážně prakticky, zatímco ve FSF je silněji zastoupeno ideologické hledisko. Je však možné, aby určitá licence softwaru vyhovovala požadavkům OSI i FSF a v mnoha případech taková situace opravdu nastává.

Samotné plánované vydání doplňku SIMULANT pod některou z open source licencí nespočívá v pouhé volbě vhodné licence a otevření zdrojového kódu. Aby takové otevření přineslo uživatelům určité výhody, je třeba upravit zdrojový kód do podoby, ve které bude využitelný i pro další osoby. V současnosti pracuji na překladu různých prvků zdrojového kódu, např. názvů procedur, funkcí a proměnných do angličtiny a opatřuji kód novými komentáři.

Jsem si přitom vědom, že pro většinu absolventů ekonomických fakult je bezplatná dostupnost nástrojů podstatnější vlastností než otevřenost zdrojového kódu. Otevřenost zdrojového kódu je však vlastností, která je podstatná především pro vývojáře, umožňuje zapojení osob napříč různými organizacemi a podporuje rozvoj jednotlivých nástrojů.

Protože diskrétní simulace za použití nástrojů s grafickým uživatelským rozhraním není open source softwarem dostatečně pokryta, předpokládám, že uvolnění nového nástroje může motivovat další osoby k zapojení do jeho dalšího rozvoje, což znamená zrychlení vývoje, efektivnější odhalování chyb a zvyšování celkové použitelnosti. Důsledkem je tedy i potenciální výhoda pro uživatele, kteří nemají žádný přímý zájem na samotné otevřenosti zdrojového kódu.

Závěr

Doplňek SIMULANT pro MS Excel byl nejprve vytvořen převážně jako výukový nástroj. Jeho další vývoj však směřuje k větším možnostem praktického využití. Tyto změny jsou motivovány především malou vybaveností českých firem komerčními produkty pro diskrétní simulaci a nedostatečnou nabídkou volně dostupných simulačních nástrojů s grafickým uživatelským rozhraním.

Cílem dalšího vývoje je vytvořit nástroj umožňující pracovat se simulačními modely způsobem, který si absolventi ekonomických oborů osvojili během základních kurzů simulačního modelování. Tohoto cíle chci dosáhnout především rozšířením doplňku SIMULANT o nové možnosti popsané v předchozí kapitole a celkovým zvýšením výkonu. K dosažení cíle by rovněž mělo přispět vydání doplňku jako open source, protože otevřenost zdrojového kódu v mnoha případech podporuje a urychluje vývoj softwaru.

Literatura

BANKS, J., 1998. *Handbook of Simulation: Principles, Methodology, Advances, Applications and Practice*. New York: John Wiley & Sons, ISBN 0-471-13403-1

COIN-OR, 2012. *The Computational Infrastructure for Operations Research*. [online]. [cit. 2014-01-29]. Dostupné z: <http://www.coin-or.org/>

DLOUHÝ, M., FÁBRY, J., KUNCOVÁ, M., HLADÍK, T., 2007. *Simulace podnikových procesů*. 1. vyd. Brno: Computer Press, ISBN 978-80-251-1649-4.

CHARVÁT, Karel, 2009. *Diskrétní simulace v MS Excel*. Praha. Diplomová práce. Vysoká škola ekonomická v Praze, Fakulta informatiky a statistika, Katedra ekonometrie.

List of discrete event simulation software, 2014. In: *Wikipedia, The Free Encyclopedia* [online]. last modified on 24 January 2014 [cit. 2014-01-29]. Dostupné z: http://en.wikipedia.org/wiki/List_of_discrete_event_simulation_software

Open-source software, 2014. In: *Wikipedia, The Free Encyclopedia* [online]. last modified on 27 January 2014 [cit. 2014-01-29] Dostupné z: http://en.wikipedia.org/wiki/Open-source_software

SIMUL8 Corporation, 2014. SIMUL8 Products. *Simul8* [online]. [cit. 2014-01-29]. Dostupné z: <http://www.simul8.com/products/index.htm>

SWAIN J. J., 2013. Simulation Software Survey. In: Institute for Operations Research and the Management Sciences. *Informs*. [online]. [cit. 2014-01-27]. Dostupné z: <http://www.orms-today.org/surveys/Simulation/Simulation.html>

JEL Classification: C44, C63, C88

Summary

Development of tool for discrete event simulation

The paper is focused on program for discrete event simulation called SIMULANT. It is a free add-in for MS Excel created in Visual Basic for Applications (VBA). The first version was developed in the context of my thesis. Further development continues. The reason for continuing development is currently mainly low availability of commercial simulation software in Czech companies and lack of user-friendly freeware or open source software for discrete event simulation.

Presentation of the SIMULANT add-in in this paper provides basic information about its capabilities, user interface, available elements, the parameters of individual elements and operation of this add-in. Part of the description of the add-in is also demonstration of recording events during simulation, which is used to generate results automatically.

Further development of SIMULANT add-in includes ongoing and planned addition of new features connected with the allocation of resources, the variability of certain model parameters in time and the control of movement of entities through the model. The aim is also to increase the overall performance of the program and to publish the add-in as open source.

Key words: discrete event simulation, SIMULANT, software, open source, Excel.

Fuzzy approach for portfolio selection problem

Maria Kobzareva

maria.kobzareva@vse.cz

Ph.D. student, Econometrics and Operations Research

Supervisor: prof. RNDr. Jan Pelikán, CSc. (jan.pelikan@vse.cz)

Abstract: One of the modern approaches to uncertainty modeling is fuzzy approach, which deals with fuzzy sets defined by fuzzy membership functions. Fuzzy approach is a modeling tool, which solves problems in which data are uncertain. This problem is frequently solved problem in many fields and it is more appropriate for real-life applications, because in practice data are hardly deterministic, but tend to change unexpectedly. The paper introduces fuzzy decomposition methodology and ranking function approach to solve fuzzy defined problems and suggests a possibility of solving portfolio optimization problem with uncertain data set using presented methods. Along with the binary portfolio optimization problem the paper suggests a modification of mentioned methods to solve integer portfolio optimization problem based on risk diversification principle. Those applications are described on a case study and provided with mathematical models and its detailed description. The paper also presents computation experiments and its results.

Key words: fuzzy variables, fuzzy linear programming, bound and decomposition method, data uncertainty, portfolio optimization problem, fuzzy number, triangle fuzzy number, ranking function

Introduction

One of the recent approaches to uncertainty modeling is fuzzy set approach, where set is described by fuzzy membership function. We distinguish two main classes of fuzzy problems: problems with fuzzy variables and problems with fuzzy parameters. The problems in which both above mentioned elements are applied are called fully fuzzified linear programming (FFLP) problems. Fully fuzzy problem is therefore generalized version of fuzzy linear programming problem where all decision parameters and variables are fuzzy numbers.

This paper presents portfolio problem formulated as fully fuzzy problem and two approaches to solve the problem. Bound and decomposition method and ranking function method are introduced to find an optimal fuzzy solution for fully fuzzy linear programming (FFLP) portfolio optimization problem. In the bound and decomposition method, the given FFLP problem is decomposed into three crisp linear programming (CLP) problems with bounded variables constraints, the three CLP problems are solved separately and by using its optimal solutions, the fuzzy optimal solution to the given FFLP problem is obtained. Fuzzy ranking function method creates a new model using ranking function for triangle fuzzy numbers and by solving the modified model fuzzy optimal solution is obtained. The methods are illustrated on a case study, results are provided.

The paper is organized as follows: in section 2 basic formulation of portfolio optimization model is introduced. Section 3 describes the fully fuzzy linear programming portfolio problem. Section 4 is associated with bound and decomposition method, in section 5 ranking function method and the properties of the presented linear ranking function with related theorems are introduced. Section 6 relates to the integer portfolio selection problem and its fuzzy modification for mentioned methods. In section 7 fully fuzzy portfolio optimization problem is solved using the above mentioned methods, results are presented and compared.

Portfolio selection problem and its mathematical formulation

Let me introduce portfolio optimization problem, where we have to select shares from a given number of shares, respecting the condition that the value of selected shares will not exceed given investment limit K . The aim is to maximize future value of selected shares, future values of shares are given. The problem is that the real share value may be less than predicted value. The purpose of the case study is to find so called fuzzy optimal solution under given circumstances, which respects unexpected decrease in future share value.

Let us define n as a given number of shares from which we should select a subset, K the limit of total expenses for selected shares, for share i a_i is the price of share, c_i is expected profit of share i , d_i is possible decrease in share profit. Binary variable x_i defines the decision to choose share i , when $x_i=1$, or decision not to choose the share in case $x_i=0$.

The basic model of choosing the shares is described below:

$$\max \sum_{i=1}^n c_i x_i, \text{ s.t. } \sum_{i=1}^n a_i x_i \leq K, x_i \in \{0,1\} i = 1,2,\dots,n. \quad (1)$$

According to this model let's create a fully fuzzy model in the following section.

Fully fuzzy linear programming portfolio optimization problem

Consider the following fully fuzzy linear programming portfolio optimization problem with n fuzzy variables, formulated as follows:

$$\max \sum_{i=1}^n \tilde{c}_i \tilde{x}_i, \text{ s.t. } \sum_{i=1}^n \tilde{a}_i \tilde{x}_i \leq \tilde{K}, \tilde{x}_i \in \{0,1\} i = 1,2,\dots,n. \quad (2)$$

Let the parameters $\tilde{a}_i$, $\tilde{c}_i$, $\tilde{x}_i$ and $\tilde{K}$ be the triangular fuzzy number (a_i', a_i, a_i'') , (c_i', c_i, c_i'') , (x_i', x_i, x_i'') and (K', K, K'') respectively. Then, the problem (2) can be written as follows [1, 249]:

$$\begin{aligned} \max (z_1, z_2, z_3) &\approx \sum_{i=1}^n (c_i', c_i, c_i'') * (x_i', x_i, x_i''), \text{ s.t. } \sum_{i=1}^n (a_i', a_i, a_i'') * (x_i', x_i, x_i'') \leq (K', K, K''), \\ (x_i', x_i, x_i'') &\in \{0,1\} i = 1,2,\dots,n. \end{aligned} \quad (3)$$

Now, since (x_i', x_i, x_i'') is a triangular fuzzy number, we can obtain bounded constraint as follows:

$$x_i' \leq x_i \leq x_i'', i = 1,2,\dots,n. \quad (4)$$

Using the arithmetic operations and partial ordering relations let's decompose the given FLPP as follows [1, 249]:

$$\begin{aligned} \max z_1 &\approx \sum_{i=1}^n \text{lower value of } ((c_i', c_i, c_i'') * (x_i', x_i, x_i'')) \\ \max z_2 &\approx \sum_{i=1}^n \text{middle value of } ((c_i', c_i, c_i'') * (x_i', x_i, x_i'')) \\ \max z_3 &\approx \sum_{i=1}^n \text{upper value of } ((c_i', c_i, c_i'') * (x_i', x_i, x_i'')) \end{aligned} \quad (5)$$

Subject to

$$\begin{aligned} \sum_{i=1}^n \text{lower value of } ((a_i', a_i, a_i'') * (x_i', x_i, x_i'')) &\leq K', \text{ for } i = 1,2,\dots,n; \\ \sum_{i=1}^n \text{middle value of } ((a_i', a_i, a_i'') * (x_i', x_i, x_i'')) &\leq K, \text{ for } i = 1,2,\dots,n; \\ \sum_{i=1}^n \text{upper value of } ((a_i', a_i, a_i'') * (x_i', x_i, x_i'')) &\leq K'', \text{ for } i = 1,2,\dots,n; \end{aligned}$$

and all decision variables are non-negative.

From the above decomposition problem, we construct the following CLP problems, namely middle level problem (MLP), upper level problem (ULP) and lower level problem (LLP) as follows:

(MLP)

$$\max z_2 = \sum_{i=1}^n c_i * x_i \quad (6)$$

subject to

$$\sum_{i=1}^n a_i * x_i \leq K, \text{ for } i = 1, 2, \dots, n; x_i \in \{0, 1\} \text{ } i = 1, 2, \dots, n.$$

At least one decision variable of the (MLP) should occur in the model and all decision variables should be non-negative.

(ULP)

$$\max z_3 = \sum_{i=1}^n c_i'' * x_i'' \quad (7)$$

subject to

$$\sum_{i=1}^n c_i'' * x_i'' \geq z_2^* ;$$

$$\sum_{i=1}^n a_i'' * x_i'' \leq K'', \text{ for } i = 1, 2, \dots, n; x_i'' \in \{0, 1\} \text{ } i = 1, 2, \dots, n.$$

In decomposition (ULP) problem at least one decision variable of the (ULP) should occur that is not used in (MLP); all variables in the constraints and objective function in (ULP) must satisfy the bounded constraints; replacing all values of the decision variables which are obtained in (MLP) and all decision variables are non-negative, where z_2^* is the optimal objective value of (MLP).

and **(LLP)**

$$\max z_1 = \sum_{i=1}^n c_i' * x_i' \quad (8)$$

subject to

$$\sum_{i=1}^n c_i' * x_i' \leq z_2^* ;$$

$$\sum_{i=1}^n a_i' * x_i' \leq K', \text{ for } i = 1, 2, \dots, n; x_i' \in \{0, 1\} \text{ } i = 1, 2, \dots, n.$$

Again, at least one decision variable of the (LLP) should occur which is not used in (MLP) and (ULP); all variables in the constraints and objective function in (LLP) must satisfy the bounded constraints; replacing all values of the decision variables which are obtained in the (MLP) and (ULP) and all decision variables are non-negative, where z_2^* is the optimal objective value of (MLP).

Remark 2.1: In the case of LP problem involving trapezoidal fuzzy numbers and variables the technique would be to decompose it into four CLP problems and then, we solve the middle level problems (second and third problems) first. Then, we solve the upper level and lower level problems and then obtain the fuzzy optimal solution to the given FLP problem involving trapezoidal fuzzy numbers and variables.

Bound and decomposition method

This section describes bound and decomposition method for solving a FFLP problem presented in [1, 250]. The algorithms to solve fuzzy portfolio optimization model with bound and decomposition method proceeds as follows.

Step 1: Construct (MLP) (6), (ULP) (7) and (LLP) (8) problems from the given FFLP portfolio optimization problem.

Step 2: Using existing linear programming technique, solve the (MLP) problem, then the (ULP) problem and then the (LLP) problem in the order only and obtain the values of all real decision variables x_i' , x_i and x_i'' and values of all objectives z_1 , z_2 and z_3 . Let the decision variables values be x_i^* , x_i^* , and x_i^{**} $i = 1, 2, \dots, n$ and objective values be z_1^* , z_2^* and z_3^* .

Step 3: An optimal fuzzy solution to the given FFLP problems is $\tilde{x}_i^* = (x_i^*, x_i^*, x_i^{**})$, $i=1, 2, \dots, n$ and the maximum fuzzy objective is $\tilde{z}^* = (z_1^*, z_2^*, z_3^*)$. (by the Theorem 3.1.).

Now, let me present the proof of the following theorem which is used in the proposed method.

Theorem 3.1 [1, 250]: Let $[x_M^*] = \{x_i^*, x_i^* \in M\}$ be an optimal solution of (MLP), $[x_U^*] = \{x_i^*, x_i^* \in U\}$ be an optimal solution of (ULP) and $[x_L^*] = \{x_i^*, x_i^* \in L\}$ be an optimal solution of (LLP) where L , M and U are sets of decision variables in the (LLP), (MLP) and (ULP) respectively. Then $\{\tilde{x}_i^* = (x_i', x_i, x_i''), i=1, 2, \dots, n\}$ is an optimal fuzzy solution to the given problem (2) where each one of x_i' , x_i , x_i'' is an element of L , M and U .

Proof [1, 250]: Let $[\tilde{y}_i] = \{\tilde{y}_i, i=1,2,\dots,n\}$ be a feasible solution of (2). Clearly, $[y_M]$, $[y_U]$ and $[y_L]$ are feasible solutions of (MLP), (ULP) and (LLP) respectively. Now, since $[x_M^*]$, $[x_U^*]$ and $[x_L^*]$ are optimal solutions of (MLP), (ULP) and (LLP) respectively, we have $Z_1([x_L^*]) \geq Z_1([y_L])$; $Z_2([x_M^*]) \geq Z_2([y_M])$ and $Z_3([x_U^*]) \geq Z_3([y_U])$.

This implies that $Z([\tilde{x}_i^*]) \geq Z([\tilde{y}_i])$, for all feasible solution of the problem (2).

Therefore, $\{\tilde{x}_i^* = (x_i', x_i, x_i''), i=1, 2, \dots, n\}$ is an optimal fuzzy solution to the given problem (2) where each one of x_i' , x_i and x_i'' , $i=1, 2, \dots, n$ is an element of L , M and U .

Hence the theorem.

Ranking function method

An efficient approach for ordering the elements is to define a ranking function $D: F(\mathfrak{R}) \rightarrow \mathfrak{R}$ which maps for each fuzzy number into the real line, where a natural order exists. Let's define orders on by [2, 24]:

$\tilde{A} \succcurlyeq \tilde{B}$ if and only if $D(\tilde{A}) \geq D(\tilde{B})$,

$\tilde{A} \preccurlyeq \tilde{B}$ if and only if $D(\tilde{A}) \leq D(\tilde{B})$,

$\tilde{A} = \tilde{B}$ if and only if $D(\tilde{A}) = D(\tilde{B})$,

where $\tilde{A}$, $\tilde{B}$ are in $F(\mathfrak{R})$. Also we write $\tilde{A} \succcurlyeq \tilde{B}$ if and only if $-\tilde{A} \preccurlyeq -\tilde{B}$. The following lemma is now immediate.

Lemma 4.1 [2, 24]. Let D be any linear ranking function then:

$\tilde{A} \succcurlyeq \tilde{B}$ if $\tilde{A} - \tilde{B} \succcurlyeq 0$ if $-\tilde{B} \succcurlyeq -\tilde{A}$,

$\tilde{A} \succcurlyeq \tilde{B}$ and $\tilde{C} \succcurlyeq \tilde{D}$ then $\tilde{A} \oplus \tilde{C} \succcurlyeq \tilde{B} \oplus \tilde{D}$.

Let's now restrict our attention to linear ranking function, that is a ranking function D such that for any $\tilde{A}$, $\tilde{B}$ belonging to $F(\mathfrak{R})$ and any $k \in \mathfrak{R}$ for any triangular fuzzy number let's use ranking function as follows [2, 24]:

$$D(\tilde{A}) = A + 1/4(A'' - A') \quad (9)$$

Then, for triangular fuzzy number $\tilde{A}$ and $\tilde{B}$, we have:

$\tilde{A} \succcurlyeq \tilde{B}$ if and only if $A + 1/4(A'' - A') \geq B + 1/4(B'' - B')$

Definition 4.1 [2, 24]. Any solution which satisfies the set of constraints of FFLP is called a feasible solution. Let S be the set of all feasible solutions of FFLP. Let's say that x^* is optimal feasible solution for FFLP if for all $x \in S$: $\tilde{c} \otimes \tilde{x}^* \geq \tilde{c} \otimes \tilde{x}$

Definition 4.2 [2, 24]. If $\tilde{A}, \tilde{b}, \tilde{c} \in \text{FT}(\mathfrak{R})$ then FFLP is transformed into the triangular fully fuzzy linear programming (TFFLP) problem.

Lemma 4.2 [2, 24]. If $\tilde{x} \in \text{FT}(\mathfrak{R})$ was positive then $D(\tilde{x}) \geq 0$.

Definition 4.3 [2, 24]. Let's define the portfolio optimization problem as follow:

$$\max \sum CX, \text{ s.t. } \sum AX \leq K, X \in \{0,1\},$$

$$\text{Where: } A = \begin{bmatrix} A_n \\ e \end{bmatrix}, \quad (10)$$

$$\text{Where: } X = X_n, K = K_1,$$

$$A_n = [A_n + 1/4A''_n - 1/4A'_n, -1/4A_n, 1/4A_n],$$

$$e = [1, -1/4, 1/4],$$

$$K_1 = K + 1/4(K'' - K'),$$

$$C_n = [C_n - 1/4C'_n + 1/4C''_n, -1/4C_n, 1/4C_n] \text{ if } \tilde{C}_n \geq 0 \text{ in the equation,}$$

$$C_n = [C_n - 1/4C'_n + 1/4C''_n, 1/4C_n, -1/4C_n] \text{ if } \tilde{C}_n < 0 \text{ in the equation.}$$

Theorem 4.1 [2, 24]. TFFLP (10) and definition (2) are equivalent.

Let S_1 and S_2 be the set of all feasible solutions of TFFLP and definition (2) respectively. Then $\tilde{X} \in S_1$ if:

Solving fully fuzzy linear programming problem where $\tilde{a}_i, \tilde{K}, \tilde{c}_i \in \text{FT}(\mathfrak{R})$ and if :

$$(AX, A'X + AX', AX'' + XA'') \leq \tilde{K}$$

$$(X, X', X'') \geq 0, X, X', X'' \in \{0,1\}$$

If:

$$AX + 1/4(AX'' + XA'' - A'X - AX') \leq K + 1/4(K'' - K')$$

$$X + 1/4(X'' - X') \geq 0$$

If:

$$X, X', X'' \in S_2$$

Hence $S_1 \subset S_2$ and $S_2 \subset S_1$.

Now we prove that any optimal solution in LP definition (2) will make an optimal solution in TFFLP (10). If X^* is optimal solution then for all $X \in S_2$ we have:

$$CX^* \geq CX$$

But if any $X = (x, x', x'')$ is assumed as fuzzy number, according to the definition of C we have:

$$D(\tilde{C} \otimes \tilde{X}^*) \geq D(\tilde{C} \otimes \tilde{X}).$$

Then:

$$\tilde{C} \otimes \tilde{X}^* \geq \tilde{C} \otimes \tilde{X}.$$

And this means that $\tilde{X}^* = (x^*, x'^*, x''^*)$ is an optimal solution for S_1 .

Remark If z^* be optimal solution of equation (10) then we have $z^* = D(\tilde{z}^*)$ that $\tilde{z}^* = \tilde{C} \otimes \tilde{X}^*$.

Integer portfolio selection problem

Another possibility to use presented methods is to extend existing portfolio selection problem on a problem where more than one share of each kind could be selected – the integer problem. To prevent the decision maker from investing into the most profitable share we should create a new constraint which will protect us against the investment risk and assure our portfolio is diversified.

To calculate the amount of risk in this paper I'd used classic economic financial indicators, presented in [4], such as average expected profit of investment E_i and coefficient of variation V_i as a measure of comparative amount of risk.

Let's calculate the average expected profit of investment E_i according to the following formula:

$$E_i = \sum_{A=1}^k c_A p_A, \quad (11)$$

where c_A is expected profit of share i in situation A ,

p_A is probability of situation A ,

E_i is average expected profit of share i .

Then let's calculate statistical dispersion as follows:

$$\sigma_i^2 = \sum_{A=1}^k (c_A - E_i)^2 * p_A, \quad (12)$$

where σ_i^2 is statistical dispersion of share i .

And obtain standard deviation according to (13):

$$\sigma_i = \sqrt{\sigma_i^2}, \quad (13)$$

where σ_i is standard deviation of share i .

The relative measure of comparative amount of risk known as coefficient of variation is then calculated as follows:

$$V_i = \frac{\sigma_i}{E_i}, \quad (14)$$

where V_i is coefficient of variation of share i .

Let's calculate the relative measure of comparative amount of risk for every share and add the following constraint (15) to the model (2):

$$\sum_{i=1}^n V_i * x_i \leq 1,5, \quad (15)$$

where V_i is coefficient of variation for share i , $x_i \{x_i \geq 0, x_i \in Z\}$, is the number of shares i which were selected and 1,5 is acceptable amount of risk in the selected portfolio (this value can vary according to the decision maker and situation at the market).

There's also a need extend binary constraint on integer constraint for variables x_i in (2).

The constraint (15) can be added to both presented models in order to calculate corresponding integer share solution. The value of acceptable risk could be changed in order to observe the influence of risk constraint on the model.

Presented method of risk measuring is one of the ways how to protect the investor from investing into one share only. There are many other methods to calculate risk. In this paper I only use the mentioned

method, as it was not the purpose of the study, but in future works it would be a good idea to compare several risk calculating methods when solving integer portfolio selection problem.

Case study

Let us solve the problem with investment limit $K=50\,000$ and set of 23 shares, from which a subset has to be selected. The data subset was obtained from web pages <http://finance.yahoo.com/>, where we used path “investment, stocks, earning dates” to collect necessary data. Each share has a current price, which was stated in the beginning of April 2013, estimated profit of the share and possible decrease in profits. The price a_i of share i , the profit c_i and decrease of the profit d_i are given in the table 1.

In the table below you can also find the highest price of share $a_{i\,high}$ and the lowest price of share $a_{i\,low}$ for a ten year period. The values a^*_i , a^*_{i-} , a^*_{i+} , c^*_i , c^*_{i-} , c^*_{i+} were calculated according to the formulas presented in the ranking function method.

Table 1: List of shares and their parameters

	a_i	c_i	d_i	$a_{i\,low}$	$a_{i\,high}$	$c_i - d_i$	$c_i + d_i$	a^*_i	a^*_{i-}	a^*_{i+}	c^*_i	c^*_{i-}	c^*_{i+}
1	2999	237	64	119	3085	173	301	3740,5	-749,8	749,8	269	-59,3	59,3
2	10400	271	31	5927	10400	240	302	11518,3	-2600	2600	286,5	-67,8	67,8
3	1633	45	15	1103	2383	30	60	1953	-408,3	408,3	52,5	-11,3	11,3
4	6825	72	14	3523	6954	58	86	7682,8	-1706,3	1706,3	79	-18	18
5	1625	30	14	998	1753	16	44	1813,8	-406,3	406,3	37	-7,5	7,5
6	1791	1	13	1791	3647	-12	14	2255	-447,8	447,8	7,5	-0,3	0,3
7	1189	10	12	768	3889	-2	22	1969,3	-297,3	297,3	16	-2,5	2,5
8	998	-13	12	830	1720	-25	-1	1220,5	-249,5	249,5	-7	3,3	-3,3
9	2253	42	11	1150	2860	31	53	2680,5	-563,3	563,3	47,5	-10,5	10,5
10	3934	44	9	1800	4476	35	53	4603	-983,5	983,5	48,5	-11	11
11	3554	62	9	2390	3554	53	71	3845	-888,5	888,5	66,5	-15,5	15,5
12	3608	61	9	457	3608	52	70	4395,8	-902	902	65,5	-15,3	15,3
13	3918	102	7	3918	5268	95	109	4255,5	-979,5	979,5	105,5	-25,5	25,5
14	4901	91	6	4038	5344	85	97	5227,5	-1225,3	1225,3	94	-22,8	22,8
15	2543	26	6	2543	6555	20	32	3546	-635,8	635,8	29	-6,5	6,5
16	768	10	5	355	940	5	15	914,3	-192	192	12,5	-2,5	2,5
17	3134	8	3	1854	3135	5	11	3454,3	-783,5	783,5	9,5	-2	2
18	1398	12	2	748	1398	10	14	1560,5	-349,5	349,5	13	-3	3
19	3299	3,7	2	2260	3299	1,7	5,7	3558,8	-824,8	824,8	4,7	-0,9	0,9
20	5272	58	2	3914	5367	56	60	5635,3	-1318	1318	59	-14,5	14,5
21	168	-1	0	100	1915	-1	-1	621,8	-42	42	-1	0,3	-0,3
22	283	-8	0	283	852	-8	-8	425,3	-70,8	70,8	-8	2	-2
23	1808	23	0	1361	1808	23	23	1919,8	-452	452	23	-5,8	5,8

I had used the algorithms described in this paper and solve the presented linear problem, results are shown in the table 2. Beside the optimal value F in the table you can also find the corresponding solution x , which denotes the share selection, the profit of presented solution without loss $\sum_{i=1}^n c_i$, potential loss of profits $\sum_{i=1}^n d_i$, and total share price $\sum_{i=1}^n a_i$.

Table 2: Steps of the algorithm and optimal fuzzy solution

method	Solution x	F	$\sum a_i$	$\sum c_i$	$\sum d_i$
MLP	1 1 1 1 1 0 0 0 1 0 1 1 1 1 0 1 0 0 0 1 0 0 1	1104	49564	1104	187
LLP	1 1 1 1 1 0 0 0 1 0 1 1 1 1 1 0 1 0 0 0 1 0 0 1	917	49564	1104	187
ULP	1 1 1 1 1 0 0 0 1 0 1 1 1 1 1 0 1 0 0 0 1 0 0 1	1291	49564	1104	187
ranking	1 1 1 0 1 0 0 0 1 0 1 1 1 1 1 0 1 0 0 0 0 0 0 1	1306	37467	974	171
	0 0		0	0	0
	1 1 1 0 1 0 0 0 1 0 1 1 1 1 1 0 1 0 1 0 0 0 0 1		38865	986	173
robust	1 1 1 1 1 0 0 0 1 0 1 1 1 1 1 0 1 0 0 0 1 0 0 1	994	49564	1104	187

As you can see in table 2 maximum objective function value for decomposition method is $F = 1104$ with the corresponding share selection $x = (1\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 1)$, for ranking function model the objective value is $F = 1306$, corresponding share selection $x = (1\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 1)$.

As you can see from the table above, optimal solution for presented approaches is different. Using decomposition method we would select shares 1 to 5, 9, 11 to 14, 16, 20 and 23, total expected profit would be 1104, on the other hand if we use ranking function method we would obtain the following solution – shares 1 to 3, 5, 9, 11 to 14, 16, 18 and 23 would be selected with total expected profit 1306. If we compare those results with the solution, obtained from solving the same problem with robust optimization model, presented in my recent paper [3], we would find out that robust solution corresponds with decomposition method solution, though with a lower value of objective function.

The disadvantage of solution, obtained by both presented methods is the fact, that we obtain the solution x as a triangle fuzzy number, which doesn't give us the idea, which shares to select, as there is always three values for each share. In the presented case study the three solutions for bound and decomposition method are the same, and in this particular situation we know which shares should be selected, but that will not always be the case.

To avoid the problem, which was mentioned above in ranking function method we can use the simplified version of the model, in which we assume, that the only parameter, presented by a triangle fuzzy number is the future profit of the shares, as it is the only parameter that is not known at the moment we choose to buy the shares. In this model we use ranking function to obtain the future profit value and then solve the classic portfolio problem (2), using ranking function coefficients for the share profit values. Another possibility is to use ranking function to obtain ranking prices in order to select the shares according to possible future changes in share prices along with ranking profit values in one model. In this paper I'd try both approaches, you can find the results of computation experiments in the table below.

Table 3: Optimal solution for simplified ranking function method

method	Solution x	F	$\sum a_i$	$\sum c_i$	$\sum d_i$
ranking a_i, c_i	1 1 1 0 1 0 0 0 1 1 1 1 1 1 1 0 0 0 0 0 1 0 0 1	1155	45905	1066	177
ranking c_i	1 1 1 1 1 0 1 0 1 1 1 1 1 1 1 0 1 0 0 0 0 0 0 1	1203	49415	1100	206

As you can see from the table above, the ranking function method suggests a possibility of share selection, where the first row also takes into account that the share price could be different, in which case total amount, invested in the shares is less than 46 000, the expected profit is 1066, which is only 38 points less comparing to the fully fuzzy models values, and potential loss is 177. The second solution takes into account changes only in share profit value, in which case we should select shares 1 to 5, 7, 9 to 14, 16 and 23 and total expected profit would be 1203, real profit for the day of purchase 1100 and potential loss 206 points.

To get more diverse results for presented methods I'd simulated a problem with investment limit 100 000 and 300 shares which values were obtained based on ranges from real share set. I had run the presented

techniques on the created model. You can find the corresponding results in table 4, where n is the number of selected shares.

Table 4: Optimal fuzzy solution for simulated model with 300 shares

method	n	F	$\sum a_i$	$\sum c_i$	$\sum d_i$
MLP	54	8298	99660	8298	1724
LLP	53	6610	99457	8283	1673
ULP	54	10022	99660	8298	1724
Ranking c_i	58	9207	99770	8217	1979

It's interesting that the best solution might be the one obtained by LLP, where the expected profit at the day of the purchase is only 15 points less (one share less) than in MLP and ULP, but the expected loss is 69 points less, along with the 203 points saved in the invested amount. Modified ranking function method does not provide satisfying results.

Table 5 presented below contains the result for integer portfolio selection problem solved with bound and decomposition method and simplified ranking function method. The amount of acceptable risk was set to 1,5 for all the problems except for the LLP in bound and decomposition method, where it was lowered to 1,3 due to model assumptions.

Table5: Optimal fuzzy solution for integer portfolio problem

method	parameters	Solution x	F	$\sum a_i$	$\sum c_i$	$\sum d_i$
MLP	$r=1,5$ $K=50$	8 1 7 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0	2584	49741	2584	655
LLP	$r=1,3$ $K=48$	7 1 6 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0	1726	45109	2302	576
ULP	$r=1,5$ $K=60$	8 1 7 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0	3298	52134	2624	674
Ranking	c_i	8 1 7 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0	2912	49741	2584	655
Ranking	a_i, c_i	8 0 7 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0	2662	40966	2343	638

The solution presented in table 5 presents share selection for both presented techniques, corresponding investment amount, expected profit value and real profit at the day of purchase. As you can see from the table, the portfolio profit is twice as big as it was when we were able to buy each share only once, although the potential loss is also three times greater, but even though the presented integer solution is more profitable then the binary solution.

Conclusion

This paper offers a possibility to solve portfolio optimization problem with uncertain share values using two fuzzy approaches – bound and decomposition method and ranking function method. The problem takes into account the fact that profits of the shares could not be the same as predicted, but real value of share profit could be less. This paper concentrates on two fuzzy approaches for portfolio optimization model, which use defuzzification method to solve the problem. Both approaches work with a fully fuzzy linear programming problem, defined by triangle fuzzy numbers in objective function, constraints and parameters. A method to solve integer portfolio selection problem based on risk diversification using described techniques is presented. The presented methods are described on a case study, provided with computation experiments and comparison of results for both presented approaches.

References

- [1] M. JAYALAKSHMI, P. PANDIAN A New Method for Finding an Optimal Fuzzy Solution For Fully Fuzzy Linear Programming Problems. In: *International Journal of Engineering Research and Applications (IJERA)* ISSN: 2248-9622 Vol. 2, Issue 4, July-August 2012, pp.247-254
- [2] T. ALLAHVIRANLOO, F. HOSSEINZADEH LOTFI, M. KH. KIASARY, N. A. KIANI, L. ALIZADEH Solving Fully Fuzzy Linear Programming Problem by the Ranking Function. In: *Applied Mathematical Sciences*, Vol. 2, 2008, no. 1, 19 - 32

- [3] J. PELIKÁN, M. KOBZAREVA Method for Solving Robust Optimization Models, J. Pelikán, M. Kobzareva Method for Solving Robust Optimization Model. In: Mathematical Methods in Economics 2013. Jihlava: College of Polytechnics Jihlava, 2013, s. 706-711. ISBN 978-80-87035-76-4.
- [4] P. MAREK Studijní průvodce financemi podniku, Praha: Ekopress, 2009, s. 76-94. ISBN 978-80-86929-49-1

JEL Classification: C6, G1

Acknowledgements:

This research was supported by the project 403/12/1947 founded by the Grant Agency of the Czech Republic and the project F4/11/2013 founded by the Internal Grant Agency of the University of Economics, Prague.

Shrnutí

Fuzzy přístup k řešení úlohy optimalizace portfolia

Jedním ze způsobů modelování neurčitostí v optimalizačních modelech je fuzzy přístup, který pracuje s fuzzy množinami, popsanými pomocí funkce příslušnosti. Jelikož fuzzy množiny jsou nástrojem pro matematické modelování vágně popsaných procesů, které nejsou přesně ohraničené, fuzzy přístup je tak modelovacím nástrojem, který umožňuje řešit úlohy, které obsahují neurčitá data. Tento problém je rozšířen v mnoha oblastech a je více vhodný pro reálné úlohy, jelikož v praxi zřídka potkáte úlohu, kde data jsou předem pevně určena a nemají tendenci se náhodně měnit. Tento článek popisuje metodu fuzzy dekompozice a metodu funkce pořadí a navrhuje možnost řešit model optimalizace portfolia s neurčitou množinou dat pomocí předložených přístupů. Vedle bivalentní úlohy optimalizace portfolia článek také navrhuje možnost řešení celočíselné úlohy optimalizace portfolia, založenou na metodě diverzifikace rizika. Problém je popsán na případové studii, předložen ekonomický i matematický model úlohy, algoritmus řešení a ilustrace a porovnání výsledků z případové studie.

Klíčová slova: fuzzy čísla, trojúhelníková fuzzy čísla, fuzzy proměnné, metoda hranic a dekompozice, fuzzy lineární programování, neurčitá data, problém optimalizace portfolia, funkce pořadí.

Analýza vztahů exportu Ruska, ceny ropy a dalších veličin

Elena Kuchina

xkuce10@isis.vse.cz

Doktorand oboru Ekonometrie a operační výzkum

Školitel: doc. Ing. Tomáš Cahlík, CSc., (cahlik.tomas@gmail.com)

Abstrakt: Práce je zaměřena na ekonometrickou analýzu vztahů mezi celkovým exportem Ruské federace, světovou cenou ropy, nabídkou ropy ze strany jejích největších světových exportérů (země organizace OPEC) a celkovou spotřebou ropných produktů v regionu OECD Europe, určující stranu poptávky.

Klíčovou koncepcí pro zkoumání zmíněné problematiky se stal VEC model (vektorový model korekce chyb) vzhledem k nestacionaritě časových řad zkoumaných makroekonomických proměnných a k existenci kointegračních vztahů mezi nimi. Pro zjištění maximální délky zpoždění byl proveden předběžný krok ve formě konstrukce VAR modelu. Tento model byl prohlášen za uspokojivé východisko pro kointegrační analýzu po vynechání makroproměnné popisující nabídkovou stranu (množství ropy produkované organizací OPEC). Dlouhodobý vztah byl odhalen mezi logaritmy celkového exportu Ruska a světové ceny ropy, a taktéž mezi logaritmy spotřeby ropných produktů v regionu OECD Europe a světové ceny ropy, avšak mezi logaritmy celkového exportu Ruska a spotřeby ropných produktů v OECD Europe dlouhodobá dynamika prokázána nebyla.

Na závěr článku byla v souvislosti s VEC modelem zkoumána také kauzalita mezi proměnnými v Grangerově pojetí s následnou aplikací „Impulse-Response“ analýzy.

Klíčová slova: Grangerova kauzalita, kointegrační vztah, „Impulse-Response“ analýzy, VAR model, VEC model.

ÚVOD

Rusko je v současné době největším producentem ropy na světě, podle údajů z roku 2012 celkový počet vytěžené ropy činil 10,37 milionu barelů denně¹, což je o 370 tisíc barelů více než ve stejném roce činila těžba v Saudské Arábii. Z toho je intuitivně zřejmé, že mezi celkovým exportem Ruska a světovou cenou ropy existuje vztah, přičemž tento vztah může být obousměrný. Na jedné straně cena ropy má vliv na celkový export Ruska přes export ropy, na straně druhé celkový export Ruska přes export ropy může mít vliv na tvorbu ceny ropy vzhledem k topové pozici Ruska v žebříčku největších producentů ropy.

Průběh vývoje celkového exportu se velmi podobá vývoji exportu ropy. Podíl vývozu ropy na celkovém vývozu Ruska od roku 2000 do roku 2005 vykazoval mírně rostoucí trend, zatímco od roku 2005 do roku 2012 se podíl vývozu ropy na celkovém vývozu zboží z Ruska ustálil na přibližné stejné úrovni a se pochybuje kolem 35 %.

Čili z výše uvedeného hypotézu vzájemného determinujícího vlivu cen ropy a celkového exportu Rusku lze pokládat za intuitivně oprávněnou, kterou ovšem je třeba prověřit metodami statisticko-ekonometrické analýzy.

Součástí zmíněné problematiky je taktéž provázanost cen ropy a tím i celkového exportu Ruska s objemem produkce ropy organizace OPEC, sdružující 12 zemí vyvážejících ropu, z nichž 11 zemí patří mezi největší světové exportéry ropy², z čehož plyne značná role organizace OPEC jako cenotvůrce ropy. Zmíněné faktory mají vliv na export Ruska především ze strany nabídky, na straně poptávky je určujícím vlivem spotřeba ropy v jednotlivých regionech. Mezi topové importéry ropy vyvážené z Ruska patří především Německo, Nizozemí, Polsko, Čína, USA, Finsko, Francie, Švédsko, Itálie, Litva, Španělsko a Japonsko.³ Z všech dodávek ropy z Ruska do zahraničí podíl připadající na Atlantický trh činí 80 %

¹ Central Intelligence Agency, The World Factbook. *Crude oil – production*. [Online]. [cit 2013-11-17]. Dostupné z: <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2241rank.html>.

² U.S. Energy Information Administration. *Top World Oil Net Exporters, 2012*. [Online]. [cit 2013-11-25]. Dostupné z: <http://www.eia.gov/countries/index.cfm?topL=exp>.

³ U.S. Energy Information Administration. *Russia*. [Online]. [cit 2013-11-25]. Dostupné z: <http://www.eia.gov/countries/analysisbriefs/Russia/russia.pdf>.

(Evropa, Střední východ, Afrika a východní pobřeží Severní Ameriky), zbylý podíl připadá na Tichooceánský trh (např. Čína, Korea, Japonsko, Thajsko a Spojené státy).⁴

Cílem daného článku je analýza vzájemného vztahu mezi celkovým exportem Ruska, světovou cenou ropy, nabídkou ropy ze strany jejích největších světových exportérů a světovou poptávkou ropy. Zjištění charakteru vztahů mezi definovanými proměnnými tj. získání informace, jak o krátkodobé dynamice, tak i o dlouhodobém vztahu. Součástí analýzy je také zjištění, jaké proměnné předcházejí změně jiné proměnné pomocí testování Grangerovy kauzality, a taktéž analýzy intenzity reakce v jedné časové řadě vyvolané exogenním šokem v jiné pomocí „Impulse-Response“ analýzy.

Článek je strukturován do tří částí, jimiž jsou datová základna a metodika postupu, určující návaznost kroků, analýza průběhu časových řad výše zmíněných proměnných a část empirická, popisující jednotlivé kroky provedené analýzy. Práce je zakončena shrnujícími informacemi o výsledcích provedeného výzkumu.

Datová základna a metodika postupu

Analyzovaná data

Táto část projednává o vstupních datech, na níž je aplikována staticko-ekonometrická analýza, návaznost, jejichž kroků je určena v části věnované *Metodice postupu*. Pro potřeby makroekonomického modelu byla užita měsíční data za období 01.2001-05.2013, tj. se jedná o 149 pozorování následujících makroekonomických proměnných:

- *Celkový export Ruské Federace v miliardách eur* – data byla získána z finančního informačního portálu QUOTE.RBC.RU⁵.
- *Světová cena ropy typu Brent v dolarech za barel* – zdrojem dat posloužil EIA (U.S. Energy Information Administration)⁶.
- *Celková spotřeba ropných produktů v regionu OECD Europe* v tisících barelů za den - zdrojem dat posloužil EIA (U.S. Energy Information Administration)⁷.
- *Množství ropy produkované organizací OPEC v milionech barelů za den* – zdrojem dat posloužily *Monthly Oil Market Reports* organizace OPEC⁸.

Vzhledem k tomu, že mezi dvanácti největšími importéry ropy vyvážené z Ruska osm jsou členy OECD Europe, celkovou *spotřebu ropných produktů v regionu OECD Europe* lze pokládat za faktor působící na *celkový export Ruska* ze strany *poptávky*. V roli faktoru působícího na ropný trh z *nabídkové* strany vystupuje v dané analýze *množství ropy produkované organizací OPEC*.

Metodika postupu

V této části bude definován postup, který byl aplikován na zkoumaná data v navazující empirické části. Veškeré potřebné výpočty byly provedeny v ekonometrickém softwaru EViews 6.

Analýza je započata zjištěním řadu integrace pomocí rozšířeného DF testu (Hušek, 2007, str. 130) aplikována na logaritmovaná data se zahrnutím konstanty a trendu. Následně pomocí modulu SEAT bude zkoumán způsob generování časové řady, umožňující rozkládat časové řady na její nepozorovatelné

⁴ Finance-Aid.ru. *Neftjanaja promyshlennost Rossii*. [Online]. [cit 2013-11-26] Dostupné z: <http://finance-aid.ru/finance/articles2/35-neftyanaya-promyshlennost-rossii.html>.

⁵ QUOTE.RBC.RU. *Objem Eksporta*. [Online]. [cit 2013-10-15]. Dostupné z: <http://quote.rbc.ru/macro/indicator/1/643.shtml>.

⁶ U.S. Energy Information Administration. *EUROPE Brent Spot Price*. [Online]. [cit 2013-10-25]. Dostupné z: <http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RBRT&f=M>.

⁷ U.S. Energy Information Administration. *Total Consumption of Petroleum Products*. [Online]. [cit 2013-10-25]. Dostupné z: <http://www.eia.gov/cfapps/ipdbproject/iedindex3.cfm?tid=50&pid=54&aid=2&cid=&syid=2000&eyid=2013&freq=M&unit=TBPD>

⁸ OPEC. *Monthly Oil Market Reports*. [Online]. [cit 2013-11-25]. Dostupné z: http://www.opec.org/opec_web/en/publications/338.htm.

komponenty (Gómez, Maravall, 1997, str. 57) a tím určit i řad ARIMA modelu. *Autoregresní integrované procesy klouzavých průměrů* řádu p, d, q , tj. $ARIMA(p,d,q)$ jsou integrovanými procesy řádu d , tj. $I(d)$ (Arlt, 2009, str. 40). Protože neexistuje přesný teoreticko-ekonomický model definující vzájemné interakci mezi zkoumanými proměnnými, východiskem pro další analýzu byl zvolen VEC model, umožňující rozlišit jak dlouhodobou, tak i krátkodobou dynamiku mezi zkoumanými časovými řadami. Odpovídající tvar vektorového modelu korekce chyb lze vyjádřit jako (Hušek, 2007, str. 251)

$$\begin{aligned}\Delta X_t &= \rho_1 Z_{t-1} + \delta_1 U_{t-1} + (\Delta X_t, \Delta Y_t)_{-i} + \varepsilon_{1t}, \\ \Delta Y_t &= \rho_2 Z_{t-1} + \delta_2 U_{t-1} + (\Delta X_t, \Delta Y_t)_{-i} + \varepsilon_{2t},\end{aligned}\quad (2.1)$$

kde lineární kombinace integrovaných proměnných řádu jedna je stacionární. Parametry δ_1 a δ_2 měří rychlost přizpůsobení k rovnováze v čase a měly by být záporné.

Před užitím VEC modelu je třeba provést odpovídající kointegrační analýzu mezi integrovanými proměnnými stejného řádu pro zjištění informací o dlouhodobých vztazích. K analýze kointegrací je užit Johansenův test, v rámci kterého jsou počítány dvě statistiky: *Trace* statistika a *Max-eig* statistika. Rozlišuje se několik případů specifikace deterministického trendu:⁹

- Úrovňová data neobsahují deterministický trend, kointegrační rovnice nemá konstantu.
- Úrovňová data neobsahují deterministický trend, kointegrační rovnice má konstantu.
- Úrovňová data obsahují deterministický trend, ale kointegrační rovnice má pouze konstantu.
- Úrovňová data a kointegrační rovnice obsahují lineární trend.
- Úrovňová data obsahují kvadratický trend a kointegrační rovnice má lineární trend.

O výběru konkrétního způsobu bude podrobněji projednáno v empirické části.

Vstupní informaci jak pro kointegrační analýzu, tak i pro odhad VEC modelu je *počet zpoždění* použitých v modelu. K určení optimální délky zpoždění slouží informační kritéria, jejichž hodnoty jsou získány z neomezeného $VAR(p)$ modelu, pro který jsou počítána následující informační kritéria (Hušek, 2007, str. 242):

- Akaikeho informační kritérium – $AIC(p)$,
- Schwartzovo informační kritérium – $SBIC(p)$,
- Hannanovo-Quinnovo informační kritérium – $HQIC(p)$.

Po konstrukce $VAR(p)$, resp. $VEC(p)$ modelu a provedení potřebných diagnostických testů lze přejít ke zkoumání krátkodobé dynamiky pomocí *Grangerovy kauzality*, umožňující určit změny jaké proměnné předchází změnám jiné proměnné (Hušek, 2007, str. 245). V návaznosti na definování kauzální závislosti v Grangerově smyslu bude provedena „*Impulse-Response*“ analýza, která zkoumá, jakou reakci v jedné časové řadě vyvolá impuls v jiné časové řadě v rámci systému (Arlt, 2009, str. 177).

Průběh makroekonomických časových řad

V této kapitole bude rozebrán průběh časových řad makroekonomických veličin v podobě celkového exportu Ruska, světové ceny ropy Brent, celkové spotřeby ropných produktů v zemích OECD Europe a množství ropy produkované organizací OPEC.

Využitím logaritmické transformace se výkyvy v hodnotách jednotlivých proměnných zmenšily, ovšem nebyly úplně odstraněny. Největší nárůst ceny ropy byl zaznamenán v červenci v roce 2008, kdy cena ropy Brent dosáhla výši 132,72 \$/barel.¹⁰ Existuje několik úhlů pohledu na vysvětlení příčiny takového ropného šoku zaznamenaného v roce 2008. Na jedné straně působí rychle se zvětšující poptávka po ropě v důsledku ekonomického růstu u rozvíjejících se tržních ekonomik v Asii a na Středním východě, zejména v Číně a Indii. Na straně druhé nabídka ropy se nachází ve stadiu stagnace. Jedním z hlavních důvodů poklesu růstu ve výrobě ropy se zdá být dlouhotrvající doba pomalých investic do nabídky ropy a výskyt přerušení dodávek, kde nejvýznamnějšími přerušeními byly zejména v nigerijské dodávce ropy v důsledku povstaleckých útoků, stejně jako pokles venezuelské a irácké výrobní kapacity. Vzhledem

⁹ EVIEWS 6 User's Guide II, str. 364-365.

¹⁰ U.S. Energy Information Administration. *EUROPE Brent Spot Price*. [Online]. [cit 2013-11-23]. Dostupné z: <http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RBRT&f=M>.

k rostoucí nerovnováze mezi poptávkou po ropě a nabídkou ropy cena ropy začala stoupat. Současně existují i alternativní objasnění tak prudkého nárůstu ceny ropy, jako jsou zavádějící vládní politika, trendy ve vývoje makroekonomických proměnných (měnový kurz) a role finančních investorů (spekulace).¹¹ Poté v tomtéž roce v prosinci došlo k velmi prudkému poklesu ceny ropy na 39,95 \$/barel,¹² v důsledku recese, a jako následku klesající poptávky po ropě.¹³

Problém takových výkyvů může být vyřešen zavedením dummy proměnných, pokud nutnost jejich začlenění bude prokázána nesplněním klasických požadavků reziduí výsledného modelu.

Vzhledem k tomu, že pro kvalitu ekonometrické analýzy je stacionárnost časových řad důležitá, je třeba její existenci ověřit. Přítomnost jednotkového kořenu lze potvrdit aplikací testů jednotkových kořenů a následnou analýzou časových řad zkoumaných proměnných z hlediska jejich procesu generování.

Rozšířené Dickeye-Fullerovy (ADF) testy jednotkových kořenů byly aplikovány na logaritmované, ovšem sezonně neočištěná data, neboť síla testu jednotkových kořenů klesá v případě sezonně očištěných dat (Hušek, 2007, str. 131). Problém kvantifikace významné sezonnosti v modelu lze řešit, jak již bylo zmíněno výše, pomocí umělých nula-jednotkových proměnných. Přítomnost jednotkového kořenu při zahrnutí konstanty a trendu byla potvrzena u časových řad všech uvažovaných makroekonomických veličin na 1% hladině významnosti. K zjištění řádu integrace je třeba aplikovat ADF na první difference. Hypotéza jednotkového kořenu byla zamítnuta u logaritmu *světové ceny ropy* a *produkce organizace OPEC* na 1% hladině významnosti, u proměnných: *celkových export Ruska* a *spotřeba ropných produktů v evropských zemích OECD* na 10% hladině významnosti. Čili výsledkem provedené analýzy je zjištění, že makro data jsou integrována řadu 1, tj. jsou typu I(1). V následující tabulce (tabulka 1) jsou uvedeny *p*-hodnoty ADF (zahrnutí konstanty a trendu) testů pro logaritmované hodnoty proměnných a pro difference jejich logaritmovaných hodnot.

Tabulka 3: ADF testy jednotkových kořenů

	Price_oil	Export_R	Consump_OECD_E	Production_OPEC
log	0,125	0,437	0,981	0,463
dif_log	0,000	0,053	0,082	0,000

Po zjištění řádu integrace za pomoci testu jednotkového kořenu, lze přejít k analýze časových řad z hlediska jejich procesu generování, tj. ke zjištění řádu procesu ARIMA, které taktéž umožní prověřit řád integrace určený při aplikaci testu jednotkových kořenů. Táto analýza, umožňující zjištění způsobu generování časové řady, může být uskutečněna za pomoci modulu software Eviews TRAMO/SEATS vyvinutého Viktorem Gómezem a Augustínem Maravalem. Konkrétně zde byl využit modul SEATS, což je program sloužící k rozkladu časové řady na její nepozorovatelné komponenty (Gómez, Maravall, 1997, str. 57). Pomocí modulu SEATS proces generování logaritmované časové řady *celkového exportu Ruska* byl rozpoznán jako ARIMA (1, 1, 0), způsob generování logaritmované časové řady *světová cena ropy* byl odhalen jako ARIMA (0, 1, 1), v případě *spotřeby ropných produktů v evropských zemích OECD regionu* proces generování byl identifikován jako ARIMA (2, 1, 0) a nakonec logaritmovaná časová řada *množství ropy produkováné organizací OPEC* byla rozpoznána jako proces ARIMA (0, 1, 1).¹⁴ Lze tedy říci, že zjištěný způsob generování časových řad všech uvažovaných makroekonomických veličin poukázalo na stejné výsledky jako ADF testy jednotkových kořenů, tj. procesy jsou integrovány řadu jedna, tj. I(1). Toto zjištění lze považovat za klíčové, neboť slouží jako výchozí bod pro pokročilejší analýzu, kterou se zabývají navazující části tohoto článku.

¹¹ BEHR, T.: The 2008 Oil Price Shock. Competing Explanations and Policy Implications. *GPPi Global Energy Governance Project Policy Paper Series No. 1*, 2009. [Online]. [cit 2013-11-23]. Dostupné z: http://www.gppi.net/fileadmin/gppi/GPPiPP1_Oil_Prices_2009.pdf.

¹² U.S. Energy Information Administration. *EUROPE Brent Spot Price*. [Online]. [cit 2013-11-23]. Dostupné z: <http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RBRTE&f=M>

¹³ WTRG Economics. *Oil Price History and Analysis*. [Online]. [cit 2013-11-23]. Dostupné z: <http://www.wtrg.com/prices.htm>.

¹⁴ Výsledky aplikace modulu TRAMO/SEATS lze získat kontaktováním autorky.

Modelování vícerozměrných nestacionárních časových řad

V návaznosti na předchozí část, cílem této kapitoly je modelování vztahů nestacionárních časových řad s ohledem na nutnost volit specifické postupy vzhledem k riziku vzniku problému zdánlivých vztahů. Podle postupu naznačeného výše, prvním krokem je zjištění optimální délky zpoždění na základě hodnot informačních kritérií získaných z konstruovaného modelu VAR(p), jež je východiskem pro kointegrační analýzu. Dalším krokem je samotná kointegrační analýza. Nakonec po definování všech vstupních charakteristik a potvrzení základních předpokladů použití, lze aplikovat vektorový model korekce chyb VEC(p).

Zjištění optimální délky zpoždění

Optimální délku zpoždění lze stanovit za pomoci informačních kritérií, k zjištění jejich hodnot byla provedena předběžná analýza pomocí VAR modelu. Vstupními endogenními proměnnými byly zvoleny logaritmy zkoumaných makroekonomických veličin, tj. E_t , P_t , C_t a Q_t ,

kde E_t – logaritmus celkového exportu Ruska (mld. EUR), $t = 2, 3, \dots, 146$,

C_t – logaritmus spotřeby ropy v zemích OECD Europe regionu (tis. barelů/den), $t = 2, 3, \dots, 146$,

P_t – logaritmus světové ceny ropy typu Brent (dolar/barel), $t = 2, 3, \dots, 146$,

Q_t – logaritmus množství ropy produkované organizací (mil. barelů/den), $t = 2, 3, \dots, 146$

Podle hodnoty HQ informačního kritéria jako optimální se ukázala délka zpoždění rovná třem obdobím, tj. $p = 3$. Avšak při testování vlastností reziduí získaného modelu hypotéza normálního rozdělení reziduí nebyla potvrzena, ani základní $G-M$ požadavky nebyly splněny, tj. ani sériová nezávislost náhodných složek, ani jejich konstantní rozptyl nebyl prokázán. Z uvedeného vyplývá, že určený model VAR (3) takovým způsobem nelze pokládat za uspokojivý pro další kointegrační analýzu. Je zřejmé, že zkoumané časové řady jsou zatíženy problémem sezónnosti, k jehož odstranění lze použít nula-jednotkové umělé proměnné. Každé ze čtyř proměnných byly přiřazeny dvě nula-jednotkové proměnné. První proměnná nabývá hodnotu jedna v případech výkyvů dolů a v ostatních případech má hodnotu nula. Hodnota rovná jedné u druhé umělé proměnné indikuje výkyvy směrem nahoru. Po zařazení nula-jednotkových proměnných do kategorií exogenních proměnných za účelem řešení problému značné sezónnosti byl znovu aplikován model VAR(3), kde hodnoty jak HQ , tak i SC informačního kritéria indikovaly jako optimální délku zpoždění rovnou třem. Avšak po testování vlastností reziduí, nebyly výsledky znovu uspokojivé, tj. nebyla splněná ani normalita rozdělení reziduí, ani jejich sériová nezávislost.

Jako jeden z možných způsobů zlepšení vlastností odhadovaného modelu bylo vyzkoušeno vynechání jedné z endogenních proměnných. Vzhledem k tomu, že cílem tohoto článku je především analýza celkového exportu Ruska přes ropy, vynechání této proměnné ani nepřichází v úvahu. Dalším logickým krokem je ponechání proměnné P_t jako ceny komodity, přes kterou se provádí analýza celkového exportu. Při bližším pohledu na proměnné z poptávkové a nabídkové strany, tj. C_t a Q_t , si lze povšimnout, že organizace OPEC má velkou moc při vlivu na světové ceny ropy. Číselný důkaz může být shledán v hodnotě korelačních koeficientů, které zobrazuje následující tabulka (tabulka 2).

Tabulka 4: Korelační koeficienty makroekonomických proměnných (po logaritmování)

	P_t	Q_t	C_t	E_t
P_t	1,000	0,837	-0,517	0,980
Q_t	0,837	1,000	-0,219	0,846
C_t	-0,517	-0,219	1,000	-0,485
E_t	0,980	0,846	-0,485	1,000

Lze pozorovat, že největší hodnota korelačního koeficientu je právě mezi P_t a Q_t v hodnotě 0,837. Z výši uvedené tabulky si lze také povšimnout, že lineární závislost mezi E_t a Q_t je větší než mezi E_t a C_t , ovšem informace, kterou proměnná Q_t přispívá k vysvětlení proměnné E_t , může být protnutá přes proměnnou P_t . Na základě předešlé diskuze bylo rozhodnuto proměnnou Q_t vynechat z další analýzy spolu s odpovídajícími umělými nula-jednotkovými proměnnými. Na zbylé tři proměnné byl znovu aplikován VAR model, kde optimální délka zpoždění jak podle HQ , tak i podle SC informačního kritéria se ukázala,

rovná třem obdobím. Podle testů diagnostické kontroly aplikované na rezidua modelu lze považovat odhadnutý model VAR (3) z hlediska normality, autokorelace a heteroskedasticity za přijatelný¹⁵ pro navazující kointegrační analýzu.

Kointegrační analýza

K analýze kointegrací byl užit *Johansenův test*, v rámci kterého byly vypočítány dvě statistiky: *Trace* statistika a *Max-eig* statistika. Pro různé případy specifikace deterministického trendu při testování kointegrace tabulka 3 uvádí počty získaných kointegračních vektorů.

Tabulka 3: Počty kointegračních vektorů

Data Trend:	None	None	Linear	Linear	Quadratic
Test Type	No Intercept	Intercept	Intercept	Intercept	Intercept
	No Trend	No Trend	No Trend	Trend	Trend
Trace	1	1	1	2	2
Max-Eig	1	1	1	2	2

Z uvedené tabulky plyne, že hypotéza existence jednoho kointegračního vztahu se potvrdila pro první tři případy, pro další dvě situace byla akceptována hypotéza existenci dvou kointegračních vektorů. První a druhý případ může být odmítnout již na začátku, neboť jejich použití podmiňuje předpoklad nepřítomnosti žádného trendu v časových řadách, navíc první případ může být aplikován pouze, pokud střední hodnota všech časových řad je nulová, což v dané analýze neplatí. Poslední případ připouštějící přítomnost kvadratického trendu může vést k dobré shodě s daty ve výběru, ovšem může pak produkovat nepravděpodobnou předpověď vně výběru.¹⁶ Finální volba tak spočívá mezi případem připouštějícím přítomnost lineárního trendu v datech, který se ovšem v samotné kointegrační rovnici nevyskytuje a situaci, kde jak úroňová data, tak i kointegrační rovnice obsahují lineární trend.

Nejlepší způsob zahrnutí deterministického trendu byl určen podle vlastností odhadové funkce VEC modelu aplikovaného na logaritmy uvažovaných makroekonomických veličin, a byl detekován 4. případ, tj. zahrnutí lineárního trendu jak do úroňových dat, tak i do kointegračního vektoru. Dlouhodobý vztah byl odhalen mezi E_t a P_t a mezi C_t a P_t . Avšak dlouhodobá dynamika mezi E_t a C_t nebyla prokázána.

Konstrukce VEC modelu

Na základě kointegrační analýzy byly detekovány dva dlouhodobé vztahy, které vstoupí do vektorového modelu korekce chyb. Vstupními endogenními proměnnými VEC modelu jsou logaritmy makroekonomických proměnných: E_t , P_t a C_t . Problém sezónnosti byl řešen pomocí umělých nula-jednotkových proměnných vystupujících v roli exogenních proměnných. Aplikací diagnostických testů na rezidua modelu bylo prokázáno jejich normální rozdělení, sériová nezávislost¹⁷ a homoskedasticita, což umožňuje přejít ke klasickým testům statistické indukce. Koeficient vícenásobné determinace v rovnicích odpovídajících E_t a C_t indikuje o relativně dobré shodě modelu s daty, ovšem u rovnice odpovídající P_t nabývá poměrně nízkých hodnot. Protože daný model neslouží k prognózování, nýbrž pouze k analýze a vysvětlení charakteru a intenzity vztahu, přednost byla dána standardním chybám a model byl považován za uspokojivý pro další analýzu. Hodnoty t -testu indikují, že odhady parametrů korekčního chybového členu vyšly statisticky významné v případě vysvětlovaných E_t a C_t . Avšak v případě P_t ani jeden z odhadu parametrů tzv. rovnovážné chyby nebyl určen t -testem jako statisticky významný. *Rychlost přizpůsobení* k rovnováze je měřena parametry korekčního chybového členu. V rovnici odpovídající E_t návrat k ekvilibriu mezi E_t a P_t probíhá s rychlostí 0,54 (odhad parametru vyšel $-0,54$, tj. je menší než jedna a má záporné znaménko, takže problém explozivního růstu nenastává), což indikuje, že kolem 54 % nerovnováhy je opraveno v průběhu jednoho období, tj. v daném případě během měsíce. Rychlost přizpůsobení k rovnováze v čase mezi C_t a P_t určuje odhad $-0,43$, indikující procento nerovnováhy,

¹⁵Podle LM testu sériové nezávislosti autokorelace 12. řadu vyšla statisticky významná, ovšem vzhledem k vysokému řadu zpoždění v porovnání se zjištěným optimálním zpožděním, model byl považován za přijatelný.

¹⁶EVew 6 User's guide II, str. 365.

¹⁷Nulová hypotéza sériové nezávislosti ve zpoždění 12. řadu nebyla zamítnuta, ovšem na základě úvah v poznámce 15 model byl považován za přijatelný.

kteřé bude korigované během jednoho měsíce. Zajímavým zjištěním je také i skutečnost, že mezi E_t a C_t žádný přímý dlouhodobý vztah nenastává. Vzájemný vliv je uskutečněn prostřednictvím P_t . Při bližší analýze samotného *korekčního chybového členu* lze pozorovat, že v první kointegrační rovnici, popisující dlouhodobý vztah mezi E_t a P_t , trend vychází statisticky nevýznamný na rozdíl od druhé kointegrační rovnice zachycující dlouhodobou dynamiku mezi C_t a P_t , kde oprávněnost přítomnosti trendu byla potvrzena statistickou významností příslušného parametru v kointegrační rovnici.

Lze si také povšimnout, že za statisticky významné byly označeny nula-jednotkové umělé proměnné v rovnici té časové řady, u níž problém významné sezónnosti pomocí těchto proměnných byl řešen. Mezi statisticky významné odhady parametrů, vyjadřující krátkodobý vliv proměnných C_t spolu s P_t na E_t patří parametry u diferencí E_t v prvním a druhém zpoždění. V případě zkoumání krátkodobých vztahů ve druhé rovnici odpovídající C_t za statisticky významné byly označeny odhady parametrů u diferencí E_t jak v prvním, tak i ve druhém zpoždění, u C_t v prvním zpoždění a u P_t ve druhém zpoždění. A nakonec ve třetí rovnici náležející P_t koeficienty, popisující krátkodobou dynamiku, vyšly statisticky významné u diferencí P_t v prvním zpoždění a C_t ve druhém zpoždění.

Přesto, že přímá dlouhodobá dynamika mezi E_t a C_t nebyla touto analýzou prokázána, krátkodobý vztah mezi nimi zjištěn byl. A proto lze přejít k dalšímu kroku, ve kterém byla otestována kauzalita mezi proměnnými v Grangerově pojetí. Testování Grangerovy kauzality v integrovaných a kointegrovaných systémech se aplikuje na diferencované časové řady, aby byly typu $I(0)$.

Testování Grangerovy kauzality

Při testování Grangerovy kauzality se testuje nulová hypotéza, zda jedna proměnná nepodmiňuje druhou proměnnou v Grangerově pojetí, tj. zda její změny nepředcházejí změně druhé proměnné. Z výsledků testování Grangerovy kauzality plyne, že difference E_t logaritmu podmiňuje ve Grangerově pojetí difference C_t , kauzální vliv ze strany difference P_t na difference C_t nebyl na 5% hladině významnosti prokázán, avšak difference E_t a P_t společně podmiňuje v Grangerově pojetí difference C_t . Takže jejich zpožděné hodnoty přispívají ke zvětšení přesnosti předpovědi C_t , což je v souladu s VEC modelem získaným v předchozím kroku. V případě difference E_t lze pozorovat, že nulová hypotéza nebyla zamítnuta ani v případě difference C_t , ani v případě difference P_t , takže zpožděné hodnoty těchto samotných proměnných nemohou zvětšit přesnost předpovědi, a jak bylo patrné z VEC modelu, v krátkém období E_t je ovlivněn pouze svými zpožděnými hodnotami, ovšem nulová hypotéza v případě společného vlivu v Grangerově pojetí nezávislých proměnných na hodnotu závislé proměnné nemůže být na 5% hladině významnosti přijata. Nakonec při zkoumání působení proměnných v Grangerově pojetí na difference P_t , Grangerova kauzalita na 5% hladině významnosti byla prokázána v případě difference C_t , což neplatí pro případ E_t . Tak jako v předchozích situacích, společně difference proměnných: C_t a E_t podmiňují difference P_t v Grangerově pojetí, což se také shoduje s výsledky VEC modelu.

Výsledky zkoumání Grangerovy kauzality lze schematicky shrnout jako:

- $\Delta E_t \rightarrow \Delta C_t$,
- $\Delta C_t \rightarrow \Delta P_t$,

tj. difference logaritmu nabídky ropy ve Grangerově pojetí podmiňuje difference logaritmu poptávky po ropě, která předchází změně logaritmu ceny ropy.

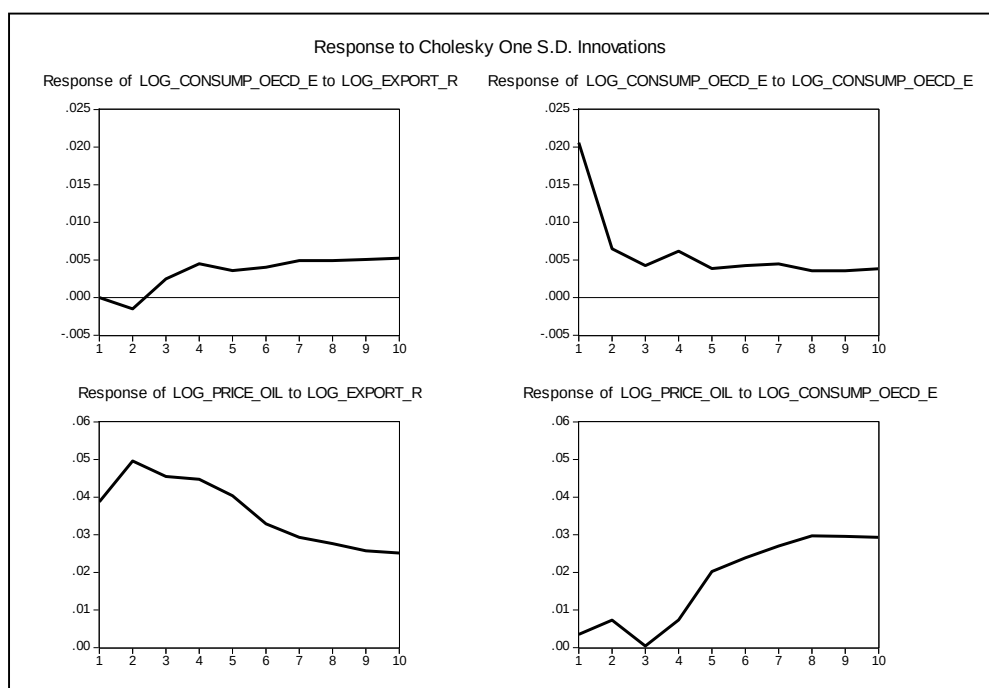
Po definování kauzální závislosti v Grangerově pojetí lze přejít k „impulse-response“ analýze umožňující zkoumat, jakou reakci v jedné časové řadě vyvolává impuls v jiné časové řadě ve vícerozměrném systému (Arl, 2009, str. 177).

„Impulse Response“ Analýza

V rámci „Impulse-Response“ analýzy byla zkoumána reakce C_t na jednotkový exogenní šok v proměnné E_t a reakce P_t na jednotkový impuls v C_t .

Pro dosažení přesné identifikace matice parametrů byla zvolena metoda Choleského dekompozice. Z grafu 1 je patrné, že proměnná C_t reaguje na šok v proměnné E_t na začátku opačným směrem do druhého období, ve kterém potom dochází k přelomu a mírnému růstu, po 4. období reakce je poměrně stejná. Jiná situace je v případě zkoumání reakce proměnné P_t na jednotkový exogenní šok v proměnné

C_t , kde po přelomu ve 3. období se reakce skokově zesiluje a poté v dalších obdobích se ustálí na stejné úrovni.



Graf 1: Impulse-Response" analýza

Závěr

Cílem článku bylo prozkoumat vzájemné vztahy mezi celkovým exportem Ruska, světovou cenou ropy, spotřebou ropné produkce v zemích OECD Europe regionu a množstvím ropy vyprodukované organizací OPEC. Vzhledem k obtížnosti definovat patřičný teoreticko-ekonomický model, popisující danou problematiku včetně definování všech endogenních a predeterminovaných proměnných, pro vzájemnou analýzu vícerozměrných časových řad, které jsou kointegrované, byl využit vektorový model korekce chyb (VEC), umožňující rozlišit mezi krátkodobou a dlouhodobou dynamikou, a také zjistit, jak rychle probíhá přizpůsobení dlouhodobé rovnováze. Ovšem pro konstrukce adekvátního modelu, splňujícího veškeré základní předpoklady, byla z analýzy vynechána proměnná odpovídající množství ropy vyprodukované organizací OPEC¹⁸, což bylo v rozporu s prvotní intuitivní představou. Ovšem dále bylo ukázané, že informace, kterou daná proměnná přispívá k vysvětlení modelu, může být promítnuta přes světovou cenu ropy (velké korelační koeficienty).

Jako dlouhodobé ekvilibrium byl determinován vztah mezi logaritmy celkového exportu Ruska a světové ceny ropy a mezi logaritmy spotřeby ropné produkce zeměmi OECD Europe regionu a také světové ceny ropy, přičemž rychlost přizpůsobení k dlouhodobé rovnováze v prvním případě byla poměrně větší. Mezi logaritmy celkového vývozu Ruska a spotřeby ropné produkce v OECD Europe regionu nebyl detekován žádný kointegrační vztah, ovšem v krátkém období pomocí Grangerovy kauzality bylo prokázáno, že změny v diferenci logaritmu celkového exportu Ruska předcházejí změnám diferenci logaritmu spotřeby ropy v OECD Europe regionu, která ovšem podmiňuje v Grangerově pojetí diferenci logaritmu světové ceny ropy, tj. difference logaritmu nabídky ropy ve Grangerově pojetí podmiňuje difference logaritmu poptávky po ropě, která předchází změně logaritmu ceny ropy. Ukázalo se, že reakce logaritmu ceny ropy na exogenní šok v logaritmu spotřeby ropy (OECD Europe) je intenzivnější, než reakce logaritmu spotřeby ropných produktů (OECD Europe) na impuls v proměnné odpovídající logaritmu celkového exportu Ruska, což je v souladu s intuitivní představou. Logaritmus samotného exportu Ruska v krátkém období je ovlivněn pouze svými hodnoty v prvním a druhém zpoždění.

¹⁸ Vzhledem k porušení podmínky normálního rozdělení reziduí.

Je zřejmé, že celkový export Ruska nemůže být determinován pouze makroekonomickými proměnnými spojenými s ropným průmyslem, neboť, jak bylo ukázáno v úvodu, podíl vývozu ropy na celkovém vývozu zboží z Ruska od roku 2005 se pochybuje kolem 35 %, což je na jedné straně dostačující podíl ke zkoumání vzájemných dlouhodobých vztahů s makroekonomickými veličinami spojených s ropným průmyslem, ovšem na straně druhé se jeví zdaleka nepostačujícím k modelování celkového exportu Ruska přes ropný průmysl. Mělo by být bráno v potaz množství dalších makroekonomických faktorů.

Literatura

ARLT, J., ARLTOVÁ, M.: *Ekonomické časové řady*. Praha: Professional Publishing, 2009. ISBN 978-80-86946-85-6.

EViews 6: *EViews 6 User's Guide II. USA*, Quantitative Micro Software, LLC, 2007.

EViews 6: *EViews 6 User's Guide I. USA*, Quantitative Micro Software, LLC, 2007.

GÓMEZ, V., MARAVALL, A.: *Programs TRAMO and SEATS. Intruction for the User*. Madrid, 1997.

HUŠEK, R.: *Ekonometrická analýza*. Praha, Oeconomica, 2007. ISBN 978-80-245-1300-3.

Internetové zdroje

BEHR, T.: *The 2008 Oil Price Shock. Competing Explanations and Policy Implications*. GPPi Global Energy Governance Project Policy Paper Series No. 1, 2009. [Online]. [cit 2013-11-23]. Dostupné z: http://www.gppi.net/fileadmin/gppi/GPPiPP1_Oil_Prices_2009.pdf.

Central Intelligence Agency. *The World Factbook. Crude oil – production*. [Online]. [cit 2013-11-17]. Dostupné z: <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2241rank.html>.

Centralnyj Bank Rossijskoj Federacii. *Eksport Rossijskoj Federacii syroj nefti za 2000-2013 gody*. [Online]. [cit 2013-11-30]. Dostupné z: http://www.cbr.ru/statistics/print.aspx?file=credit_statistics/crude_oil.htm&pid=svs&sid=vt1.

Centralnyj Bank Rossijskoj Federacii. *Vneshnjaja trgovlja Rossijskoj Federacii tovarami*. [Online]. [cit 2013-11-30]. Dostupné z: http://www.cbr.ru/statistics/print.aspx?file=credit_statistics/trade.htm&pid=svs&sid=vt.

Finance-Aid.ru. *Neftjanaja promyshlennost Rossii*. [Online]. [cit 2013-11-26] Dostupné z: <http://finance-aid.ru/finance/articles2/35-neftyanaya-promyshlennost-rossii.html>.

OPEC. *Monthly Oil Market Reports*. [Online]. [cit 2013-11-25]. Dostupné z: http://www.opec.org/opec_web/en/publications/338.htm.

QUOTE.RBC.RU. *Objem Eksporta*. [Online]. [cit 2013-10-15]. Dostupné z: <http://quote.rbc.ru/macro/indicator/1/643.shtml>.

U.S. Energy Information Administration. *EUROPE Brent Spot Price*. [Online]. [cit 2013-10-25]. Dostupné z: <http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RBRTE&f=M>.

U.S. Energy Information Administration. *Russia*. [Online]. [cit 2013-11-25]. Dostupné z: <http://www.eia.gov/countries/analysisbriefs/Russia/russia.pdf>.

U.S. Energy Information Administration. *Top Worldt Oil Net Exporters, 2012*. [Online]. [cit 2013-11-25]. Dostupné z: <http://www.eia.gov/countries/index.cfm?topL=exp>.

U.S. Energy Information Administration. *Total consumption of Petroleum Products*. [Online]. [cit 2013-10-25]. Dostupné z: <http://www.eia.gov/cfapps/ipdbproject/iedindex3.cfm?tid=50&pid=54&aid=2&cid=&syid=2000&eyid=2013&freq=M&unit=TBPD>.

WTRG Economics. *Oil Price History and Analysis*. [Online]. [cit 2013-11-23]. Dostupné z: <http://www.wtrg.com/prices.htm>.

JEL classification: C51, C52, E29, F40

Summary

Analysis of the relationships between Russian export, oil price and other variables

The aim of this article was to explore the relationship between the total Russian export, the world price of crude oil, the consumption of the petroleum products in OECD Europe and the oil OPEC production. Due to the difficulty in defining an appropriate theoretical economic model which would describe the given issue and allow to determine all the endogenous and predetermined variables, the VEC model was used to analyze the mutual relationships between the multivariate time series, which are cointegrated. Vector error correction model allows to distinguish between the short-term and long-term dynamics and also to determine how quickly the system returns to the long-term equilibrium. For the construction of an adequate model fulfilling all the basic assumptions the variable corresponding to the oil OPEC production was omitted from the model, which was contrary to the initial assumption. However, it was shown that the information by which this macro variable contributes to the explanation of the model can be projected through the world oil price (high correlation coefficients).

In this study a long-term relationship was revealed between the logarithms of the total Russian export and the world oil price, and also between the logarithms of the consumption of petroleum products in OECD Europe and the world oil price, but the long-term dynamics between the logarithms of the total Russian export and the consumption of petroleum products in OECD Europe was not proven.

However, by using the Granger causality it has been shown that in the short term the changes in the difference of the logarithm of the total Russian export precede the changes in the difference of the logarithm of the oil consumption in OECD Europe, which determines in Granger concept the difference in the logarithm of the world oil prices.

After applying "Impulse-Response" analysis it turned out that the reaction of the logarithm of oil prices on the exogenous shock in the logarithm of oil consumption (OECD Europe) is greater than the response of the logarithm of consumption of petroleum products to the impulse in the logarithm of the total Russian export, which is consistent with the intuitive idea. The logarithm of the current export of Russia in the short term is affected only by their values in the first and second lag.

It is obvious that the total Russian export cannot be determined only by macroeconomic variables associated with the oil industry, as it has been shown above. The share of the oil export in the total export of goods from Russia since 2005 is around 35%, which is on the one hand a sufficient share to explore the mutual long-term relationships between macroeconomic variables associated with the oil industry, but on the other hand it appears to be by far insufficient to model the total exports only via Russia's oil industry. Number of other macroeconomic factors should be taken into account.

Key words: Granger causality, cointegration relation, "Impulse-Response" analysis, VAR model, VEC model.

GARCH(1,1), metoda maximální virohodnosti a počet pozorování

Ing. Peter Princ
peter.princ@vse.cz

Doktorand oboru Ekonometrie a operační výzkum

Školitel: doc. Mgr. Ing. Martin Dlouhý, Dr., MSc, (dlouhy@vse.cz)

Abstrakt: Tento článek zameriava svoju pozornosť na preskúmanie vlastností modelu GARCH(1,1) z hľadiska skreslenia odhadu parametrov modelu zapríčineného malým počtom pozorovaní použitých pre odhad modelu a rozširuje súčasné vedomosti o skreslení parametrov. Pomocou Monte Carlo simulácií a analýz odhadu parametrov získaných metódou maximálnej virohodnosti je preskúmaný priestor parametrov spĺňajúci štandardné obmedzenia kladené na tieto parametre a sú kvantifikované priemerné skreslenia parametrov pre danú kombináciu ARCH a GARCH parametrov modelu, pričom sa vychádza z predpokladu nulového vplyvu konštantného člena rovnice podmieneného rozptylu na ostatné parametre modelu pri odhade pomocou metódy maximálnej virohodnosti. Zaujímavým poznatkom vyplývajúcim z tejto analýzy je rozličná skreslenosť odhadov parametrov v priestore, pre ktoré sa persistencia procesu meraná ako súčet ARCH a GARCH parametra blíži k hodnote jedna. Persistencia predstavuje rýchlosť poklesu vplyvu veľkých minulých hodnôt volatility na súčasné hodnoty podmieneného rozptylu. Pre vybrané kombinácie parametrov je preskúmaný tvar funkcie virohodnosti a je analyzovaná závislosť medzi tvarom funkcie virohodnosti, hodnotou parametrov a počtom pozorovaní zahrnutých do odhadu modelu.

Kľúčové slová: GARCH(1,1) model, metóda maximálnej virohodnosti, skreslenie parametrov, počet pozorovaní.

Úvod

Volatilita indikuje úsek časového radu, ktorý je spojený s vysokou variabilitou alebo rastúcim rozptylom. Tento fenomén hrá významnú úlohu v procese modelovania a analýzy finančných časových radov. Volatilita je kľúčovým elementom procedúr, ktoré hodnotia celkové riziko finančných aktív. Modely slúžiace na predpovedanie volatility sú používané v investičných rozhodovacích procesoch pre zachytávanie možného rizika potenciálneho investičného portfólia, v analýze VaR modelov alebo modeloch opcií. Preto nie je prekvapujúce, že v poslednom období bolo vynaloženého veľa úsilia na zostavenie modelov zachytávajúcich dynamiku volatility.

Pre finančné časové rady je charakteristické zoskupovanie volatility a leptokurtizmus. Ich ďalšou charakteristickou črtou je tzv. pákový efekt (Black, 1976). Pozorovania tohto typu vo finančných časových radoch viedli k širšiemu využívaniu modelov s podmieneným rozptylom pre odhad a predpoveď volatility. Engle (1982) navrhol vo svojej práci modelovať časovo meniaci sa podmienený rozptyl ako autoregresný podmienený heteroskedastický (ARCH) proces využívajúci oneskorené náhodné zložky. V neskorších prácach bola preukázaná nevyhnutnosť použitia vyššieho stupňa ARCH pre zachytenie dynamickej povahy podmieneného rozptylu. Všeobecný model ARCH (GARCH) navrhnutý Bollerslevom (1986) je založený na nekonečnej špecifikácii ARCH člena, ktorý redukuje počet odhadnutých parametrov z nekonečna na dva. Obidva modely ARCH a GARCH sú schopné dobre opisovať zhlukovanie volatility vo finančných časových radoch, na druhej strane zlyhávajú pri modelovaní pákového efektu. Ďalším problémom súvisiacim s GARCH modelmi je problém leptokurtizmu, t.j. pravdepodobnostné rozdelenie skúmaného procesu je špicatejšie a má tlstejšie konce ako normálne rozdelenie, ktorý nedokážu plne zachytiť. Možnosťou ako prekonať tieto problémy je použitie iného ako normálneho rozdelenia náhodných chýb.

Tento článok poukazuje na potenciálne skreslené odhady parametrov modelu GARCH(1,1), ktoré sú odhadované metódou maximálnej virohodnosti. Skreslenie odhadu parametrov finančného procesu môže byť zapríčinené malým počtom pozorovaní použitých pre odhad, zmenou hodnoty parametra procesu v dôsledku štruktúrneho zlomu alebo nesprávnym predpokladom o rozdelení náhodnej zložky procesu. V ďalšom texte sa budem venovať analýze príčin a dôsledkov skreslených odhadov parametrov modelu GARCH(1,1) s normálne rozdelenými náhodnými chybami. Pričom budeme vychádzať z teórie,

ktorú publikoval Lumsdaine (1995) vo svojom článku o vlastnostiach odhadu GARCH(1,1) modelu pomocou funkcie maximálnej vierohodnosti, kde dokázal, že zmena hodnoty konštantného člena podmieneného rozptylu nemá vplyv na zmenu odhadnutých hodnôt ARCH a GARCH parametra. Preto ďalšie analýzy sú založené na nastavenej fixnej hodnote konštantného člena podmieneného rozptylu, malom počte pozorovaní a ich spoločného vplyvu na skreslenie odhadov ARCH a GARCH parametrov.

Hlavná motivácia, ktorá ma priviedla k analýze vlastností odhadov pomocou maximálnej vierohodnosti vo finančných procesoch, tkvie v charakteristike samotných procesov, pri ktorých sme veľmi často svedkami zhľukovania (z angl. clustering) dát. Mieru zhľukovania dát v prípade GARCH(1,1) modelu predstavuje suma hodnôt ARCH a GARCH parametra. Tento jav úzko súvisí so stacionaritou samotného procesu. Ak je suma týchto dvoch parametrov rovná resp. väčšia ako jedna, tak sa modelovaný proces stáva nestacionárnym, t.j. vplyv minulých hodnôt parametrov procesu na budúce hodnoty (predikcie) podmieneného rozptylu nemá neklesajúcu tendenciu. Čím sa persistencia procesu GARCH(1,1) blíži viac k hodnote 1, tým pomalšie konverguje podmienený rozptyl k nepodmienenému rozptylu. Preto boli bližšie preskúmané rôzne prípustné kombinácie parametrov ARCH a GARCH, u ktorých je persistencia blízka hodnote 1. Zaujímavým poznatkom vyplývajúcim z tejto analýzy je rôzna skreslenosť odhadov parametrov pri rôznych preskúmaných kombináciách, kde kľúčovú rolu zohráva vyššia hodnota parametra GARCH, resp. nižšia hodnota parametra ARCH.

Hwang et al. (2004) skúmajú vplyv Bollerslevových nezáporných podmienok na parametre odhadnuté modelom GARCH(1,1). Prichádzajú k záveru, že Bollerslevové podmienky majú za následok veľké množstvo konvergenčných chýb pri odhade modelu, čoho výsledkom sú skreslené odhady parametrov. Ich analýza vychádza zo skúmania 4 rôznych kombinácií modelu; konkrétne prichádzajú k záveru, že hodnoty odhadu konštantného člena podmieneného rozptylu sú pozitívne skreslené pre malé počty pozorovaní, pre malé hodnoty ARCH parametra sú odhady ARCH parametra pozitívne skreslené a odhady GARCH parametra negatívne skreslené. S rastúcou hodnotou parametra ARCH sa stávajú odhady parametra GARCH menej skreslené. Ma et al. (2007) vyvodzujú záver, že pri malej hodnote (0,01-0,1) parametra ARCH a priemernej hodnote (0,5) parametra GARCH sú odhady parametra GARCH pozitívne skreslené. Ich práca zahŕňa hlbšiu analýzu vlastností odhadov GARCH parametra pri malej hodnote ARCH parametra na základe overenia tzv. limitnej podmienky nulovej informácie (zero-information-limit-condition), ďalej len LPNI. Táto podmienka skúma model z hľadiska slabšej identifikácie, ktorá by mohla viesť k falošnej inferencii. Následkom takejto inferencie sú odhady štandardných chýb parametrov modelu, ktoré sú veľmi vzdialené od asymptotických štandardných chýb. Samozrejme pri použití malého počtu pozorovaní na odhad parametrov modelu sa odhady štandardných chýb stávajú skreslenejšie vzhľadom na svoje skutočné asymptotické hodnoty. Avšak pri porovnaní štandardných chýb odhadnutých parametrov, ktorých hodnoty sú skreslené vzhľadom na malý počet pozorovaní, so štandardnými chybami parametrov, ktoré dosahujú skutočné hodnoty parametrov, nemožno vysloviť záver, že takto dosiahnuté odhady štandardných chýb možno považovať za skreslené.¹

Figlewski (2004) poukazuje na fakt, že príčinou skreslenia odhadu parametrov modelu GARCH(1,1) je plochý tvar logaritmizovanej vierohodnostnej funkcie v blízkosti jej optima. Preto bol tiež preskúmaný tvar logaritmizovanej vierohodnostnej funkcie pre rôzne kombinácie parametrov a rôzne počty pozorovaní. Vierohodnostná funkcia je charakteristická plochým údolovitým tvarom v blízkosti optima, ktorý sa mení so zvyšujúcim sa počtom pozorovaní alebo vplyvom zmeny hodnôt parametrov. S rastúcim počtom pozorovaní sa mení tvar (zväčšuje strmosť) funkcie v okolí optima a optimum tak leží v priestore, ktorý je optimalizačný algoritmus schopný ľahšie odhaliť. Tento fakt úzko súvisí s problematikou

¹ Pomocou MC simulácie som porovnal odhady štandardných chýb parametrov modelu GARCH(1,1) so skutočnými hodnotami štandardných chýb pre danú kombináciu parametrov, ktoré boli získané z odhadu variačno-kovariančnej matice, do ktorej boli vložené skutočné hodnoty simulovaného procesu a ktorá bola vyjadrená v analytickom tvare Hessiánu funkcie vierohodnosti. Konkrétny analytický tvar je možné nájsť napr. v Fiorentini et al. (1996). Nepublikované výsledky tejto analýzy budú poskytnuté na požiadanie.

použitia aproximácie pomocou numerického Hessiánu v optimalizačnom algoritme, ktorý v každom odhadovom kroku určuje smer, ktorým bude algoritmus postupovať v priestore parametrov².

V nasledujúcej časti definujem GARCH(1,1) model s normálne rozdelenou náhodnou zložkou, formulujem postup odhadu modelu prostredníctvom funkcie maximálnej vierohodnosti a štatistickú inferenciu vychádzajúcu z tejto metódy odhadu. V ďalšej sekcii je preskúmané skreslenie parametrov odhadnutých touto metódou a sú bližšie analyzované vybrané kombinácie parametrov. Pozornosť je venovaná taktiež preskúmaniu vlastností odhadov v prípade, ak sa persistencia procesu blíži k hodnote 1. Všetky uvedené analýzy sú študované aj z hľadiska rôzneho počtu pozorovaní použitých na odhad modelu.

Metodológia

Model GARCH (1,1)

Odhadovaným modelom bude model GARCH (1,1) navrhnutý Bollerslevom (1986), ktorý má rovnicu podmieneného priemeru definovanú:

$$w_t = \varepsilon_t$$

a rovnicu pre podmienený rozptyl:

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1},$$

kde $\omega > 0$, α a $\beta \geq 0$ a $\alpha + \beta < 1$.

Pokiaľ vychádzame z predpokladu o normalite náhodnej zložky, tj. $\varepsilon_t = \sqrt{h_t} e_t$, $e_t \sim N(0,1)$, tak potom pre výpočet parametrov modelu $\theta = (\omega, \alpha, \beta)$ použijeme metódu maximálnej vierohodnosti.

Odhad modelu pomocou metódy maximálnej vierohodnosti

Odhad parametrov modelu GARCH(1,1) prostredníctvom metódy maximálnej vierohodnosti je najčastejšie využívanou odhadovou technikou v prípade splnenia predpokladu o podmienených nezávislých a rovnako rozdelených náhodných zložkách pochádzajúcich z normálneho rozdelenia. Vierohodnostná funkcia pre T nezávislých premenných potom nadobúda tvar:

$$f(w; \theta) = \prod_{t=1}^T (2\pi\sigma_t^2)^{-\frac{1}{2}} \exp\left(-\frac{w_t^2}{2\sigma_t^2}\right)$$

a normálna logaritmizovaná vierohodnostná funkcia je:

$$l(\theta; w) = \sum_{t=1}^T -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma_t^2) - \frac{w_t^2}{2\sigma_t^2}.$$

Táto funkcia je maximalizovaná vzhľadom na parametre modelu θ . Použitie funkcie normálnej vierohodnosti pre odhady modelu má svoje opodstatnenie, pretože takéto odhady sú silne konzistentné. V prípade, ak nie je splnený predpoklad podmienenej normality náhodnej zložky, a ak sú rovnice podmieneného priemeru a podmieneného rozptylu správne špecifikované, budú odhady parametrov modelu stále konzistentné.

² V tomto článku sa nebudem bližšie venovať problematike použitia jednotlivých numerických aproximácií a analytických tvarov prvých a druhých diferencií funkcie maximálnej vierohodnosti použitých pri hľadaní optima pomocou gradientných algoritmov. Ako dobrá referencia k danej oblasti posluží napr. Fiorentini et al. (1996).

Štatistická inferencia

Odhady parametrov prostredníctvom metódy maximálnej virohodnosti sú asymptoticky normálne rozdelené, preto:

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, I^{-1}) ,$$

kde

$$I = -E \left[\frac{\partial^2 l(\theta_0; w_t)}{\partial \theta \partial \theta'} \right]$$

je záporná stredná hodnota Hessiánu. Podobne ako je to v prípade druhej derivácie, ktorá predstavuje smernicu (sklon), Hessián meria zakrivenie virohodnostnej funkcie v optime. Odhad I je realizovaný pomocou výberovej informačnej matice

$$\hat{I} = T^{-1} \sum_{t=1}^T \frac{\partial^2 l(\hat{\theta}; w_t)}{\partial \theta \partial \theta'} .$$

V prípade štandardných chýb odhadov parametrov, kde nie je splnený predpoklad podmienenej normality náhodnej zložky, sú takto odhadnuté štandardné chyby nevhodné a preto je nutné použiť robustný estimátor variačno-kovariančnej matice. V našom prípade bola použitá metóda kvázimaximálnej virohodnosti s Bollerslev-Wooldridge (1992) štandardnými chybami s ohľadom na malý počet pozorovaní použitých pre odhady modelu GARCH(1,1). Takýto estimátor dáva asymptoticky normálne odhady parametrov

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, I^{-1} J I^{-1})$$

a má tvar

$$\hat{I}^{-1} \hat{J} \hat{I}^{-1} ,$$

kde

$$\hat{J} = T^{-1} \sum_{t=1}^T \frac{\partial l(\hat{\theta}; w_t)}{\partial \theta} \frac{\partial l(\hat{\theta}; w_t)}{\partial \theta'} .$$

V prípade predpokladu o normalite $I = J$ a tým pádom sú obidva estimátory variačno-kovariančnej matice totožné.

Monte Carlo simulácie a analýza vlastností odhadov na malom počte pozorovaní

Odhady parametrov z metódy maximálnej virohodnosti sú preskúvané pomocou komplexného systému Monte Carlo (MC) simulácií. Skreslenosť odhadu parametrov modelu pri malých výberoch z finančného časového radu je študovaná pre rozsah súboru 250 a 500 pozorovaní. Hwang et al. (2004) odporúčajú použitie minimálneho počtu 1000 pozorovaní pri odhadoch GARCH modelov. Levy (2001) odporúča použitie minimálneho počtu 200-300 pozorovaní pre spoľahlivé modelovanie GARCH procesu.

Mojím zámerom je vyšetriť vlastnosti odhadov modelu pri menšom počte pozorovaní. Pri mnou zvolenom nastavení MC simulácie zodpovedá 250 pozorovaní denným údajom o finančných aktívach za posledných 12 mesiacov a 500 pozorovaní zodpovedá denným údajom v priebehu 2 rokov. Vzhľadom na fakt, že všetky predchádzajúce štúdie skúmajú problematiku skreslenia odhadov parametrov v dôsledku malého počtu pozorovaní len prostredníctvom niekoľkých možných kombinácií parametrov, zameral som sa na preskúmania možných skreslení odhadu parametrov prostredníctvom komplexnejšej analýzy. Pre tieto účely bola fixne nastavená hodnota parametra $\omega = 0,0001$ a boli preskúvané všetky možné

kombinácie parametrov α a β , pričom kombinácie boli generované prostredníctvom mriežky, kde hodnoty parametrov môžu nadobúdať hodnoty z intervalu $\langle 0,01:0,99 \rangle$ a hodnoty parametrov môžu byť zvýšené/znížené o 0,02. Tak vznikne mriežka, ktorá umožňuje komplexnejšie sledovať skreslenie ARCH a GARCH parametrov na celom intervale ich prípustných hodnôt. Podobne ako Lumsdaine (1995) a Chan et al. (2009), bolo pre každú možnú kombináciu parametrov realizovaných 500 simulácií. Následne boli pre príslušnú kombináciu parametrov a počet simulácií prevedené odhady metódou maximálnej vierohodnosti, kde boli parametre modelu ohraničené tak, ako to definoval vo svojej práci Bollerslev (1986).

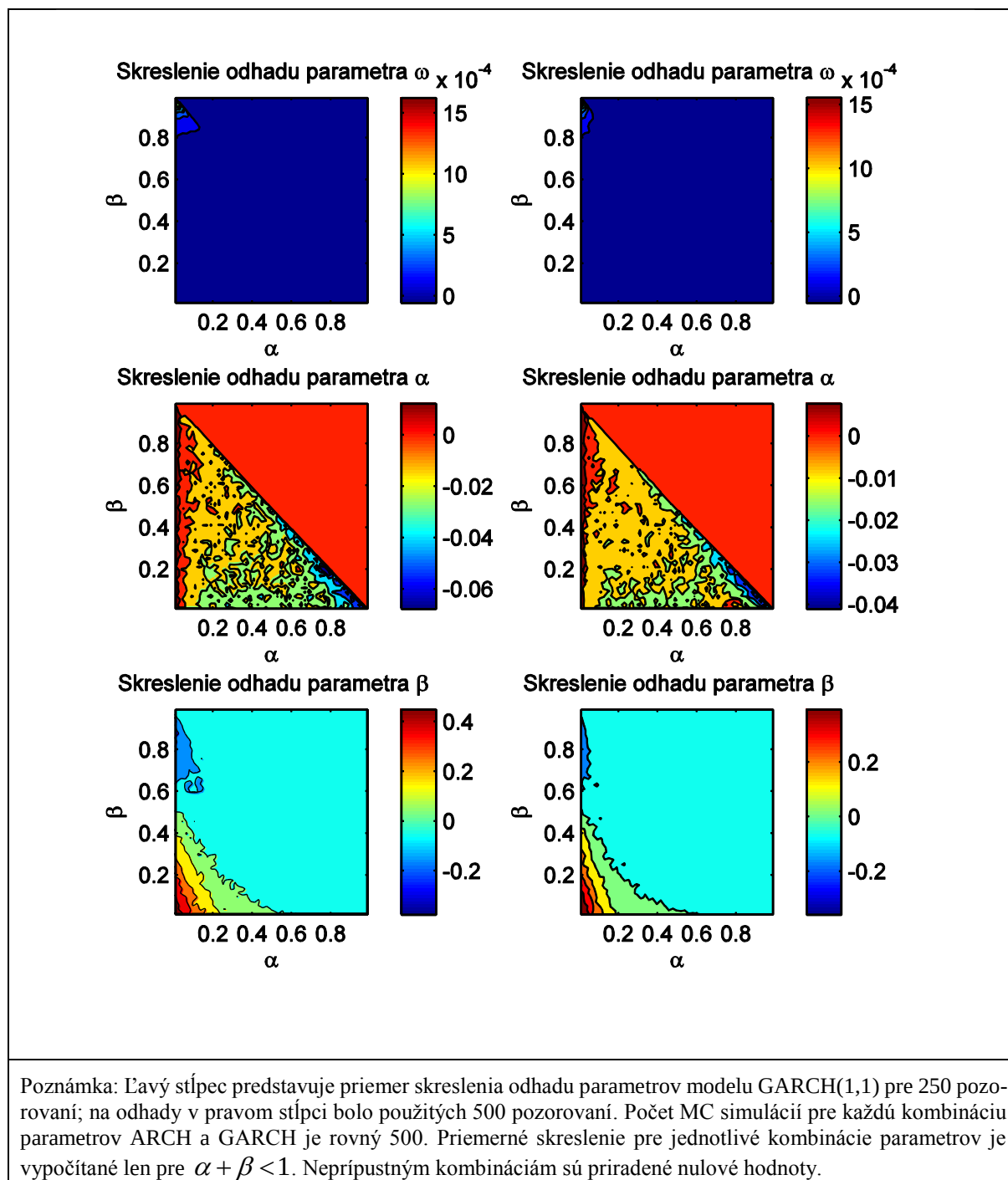
Skreslenosť odhadov GARCH(1,1) modelu

V tejto časti budú prezentované výsledky MC simulácií z realizovaných experimentov. Obrázok 1 prezentuje priemerné hodnoty skreslenia parametrov prostredníctvom tzv. vrstevnicových grafov (z angl. contour plots) pre všetky prípustné kombinácie parametrov.

Na rozdiel od Hwang et al. (2006) prevedené experimenty nepotvrdili možné skreslenie odhadov konštantného člena podmieneného rozptylu (omega) na celej prípustnej množine hodnôt parametrov. Kladné skreslenie odhadu parametra omega sa prejavuje len pri vysokých hodnotách parametra GARCH a nízkych hodnotách parametra ARCH, pričom je zjavné že toto skreslenie sa znižuje vplyvom rastúceho počtu pozorovaní. Ďalej sa mi podarilo overiť výsledky týkajúce sa pozitívneho skreslenia parametra ARCH pri malých hodnotách tohto parametra, avšak nepotvrdilo sa tvrdenie o súvisiacom negatívnom skreslení parametra GARCH. Skreslenie parametra GARCH jednak závisí od hodnoty parametra ARCH ale aj od jeho samotnej hodnoty. Zlomom v skreslení parametra GARCH je hodnota 0,5, pre nižšie hodnoty parametra GARCH je skreslenie kladné, pričom platí, že čím sa kombinácia parametrov ARCH a GARCH blíži viac k hodnotám (0,0), tým je pozitívne skreslenie väčšie. K negatívnomu skresleniu odhadu parametra GARCH prichádza v prípade ak je hodnota parametra väčšia ako 0,5 a parameter ARCH nadobúda malé hodnoty. Zároveň je nutné zdôrazniť, že v prípade veľkosti skreslenia v absolútnych hodnotách neprichádza k zmenšeniu pri prechode z 250 pozorovaní na 500 pozorovaní; prichádza len k zmenšeniu plochy na ktorej sa jednotlivé skreslenia parametrov vyskytujú.

V prípade skreslenia parametra ARCH je dôležité podotknúť, že skreslenie nedosahuje takých dramatických hodnôt ako skreslenie parametra GARCH. Záporné hodnoty skreslenia parametra ARCH súvisia najmä s blízkosťou línie pomyslenej priamky $\alpha + \beta = 1$, kde dochádza k najväčšiemu negatívnomu skresleniu odhadov. Takisto sa potvrdil predpoklad o tom, že s rastúcim počtom pozorovaní klesá absolútna hodnota skreslenia pre všetky parametre modelu GARCH(1,1). Závery vyplývajúce z takto nastavenej analýzy som bližšie preskúmal prostredníctvom analýzy tvaru a prejavu funkcie maximálnej vierohodnosti v blízkosti jej optima. Opäť budem vychádzať z poznatku o minimálnom vplyve hodnoty parametra ω na ďalšie parametre modelu GARCH(1,1).³

³ Závery vyplývajúce z analýzy Lumsdaine (1995) som podrobil hlbšiemu preskúmaniu v prípade využitia malého počtu pozorovaní pre odhady modelu GARCH(1,1). Z mojich výsledkov vyplýva, že aj v tomto prípade nemá zmena hodnoty parametra ω vplyv na schopnosť funkcie maximálnej vierohodnosti určiť optimum, t.j. zmena hodnoty parametra zníži hodnotu funkcie maximálnej vierohodnosti v novodosiahnutom optime. Preto pri použití vhodného optimalizačného algoritmu pre hľadanie parametrov modelu prostredníctvom funkcie maximálnej vierohodnosti sú vždy nájdené také odhady parametrov, ktoré sú najmenej skreslené. Odlišným prípadom je odhad parametrov modelu prostredníctvom zovšeobecnenej metódy momentov, kde zmena hodnoty parametra ω má vplyv na hodnotu kritériálnej funkcie, t.j. zníži jej hodnotu. Takýto jav je v tomto prípade nežiaduci, pretože táto metóda je založená na hľadaní minima kritériálnej funkcie.



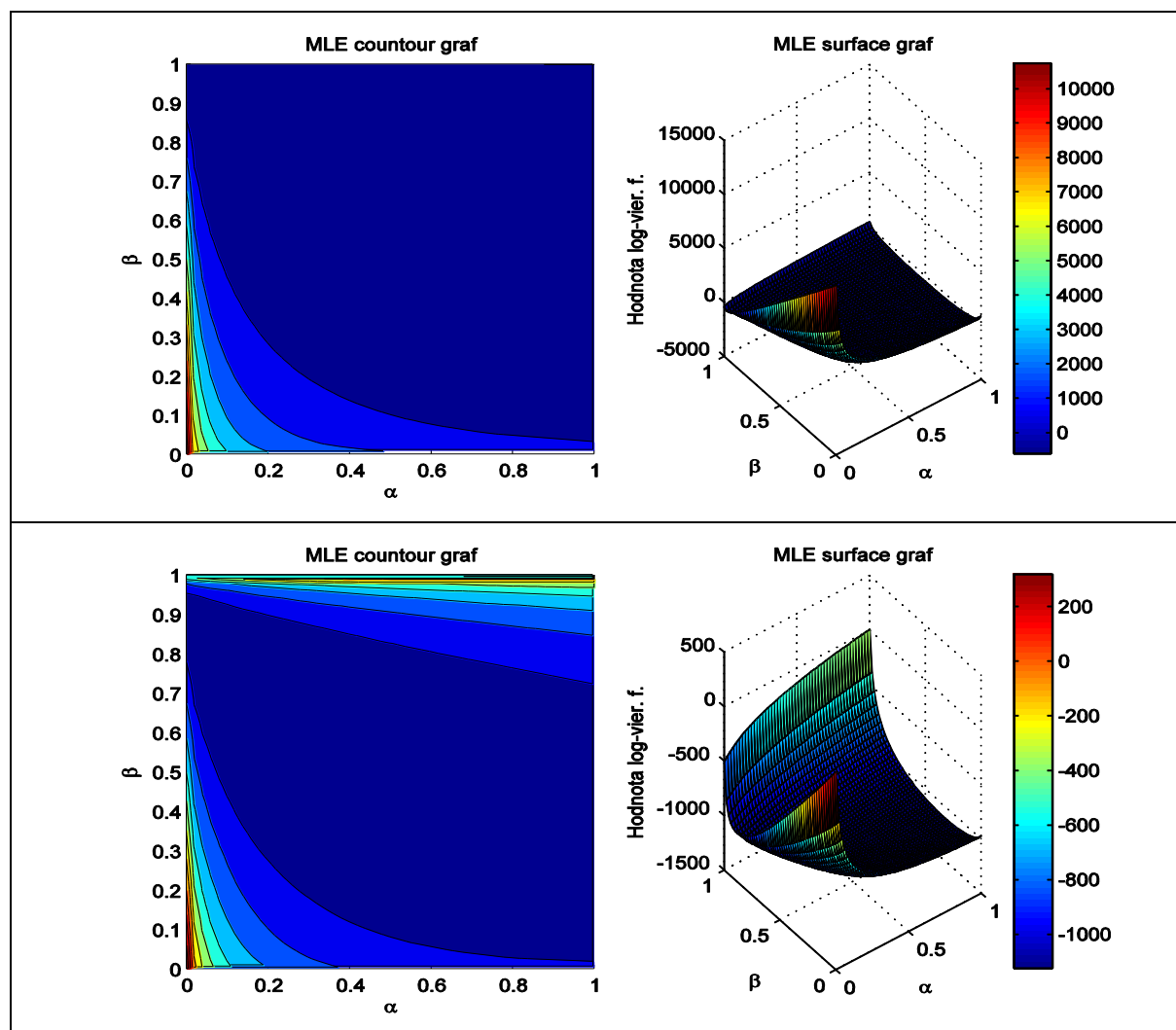
Obrázok 25: Priemerné skreslenie parametrov modelu GARCH(1,1)

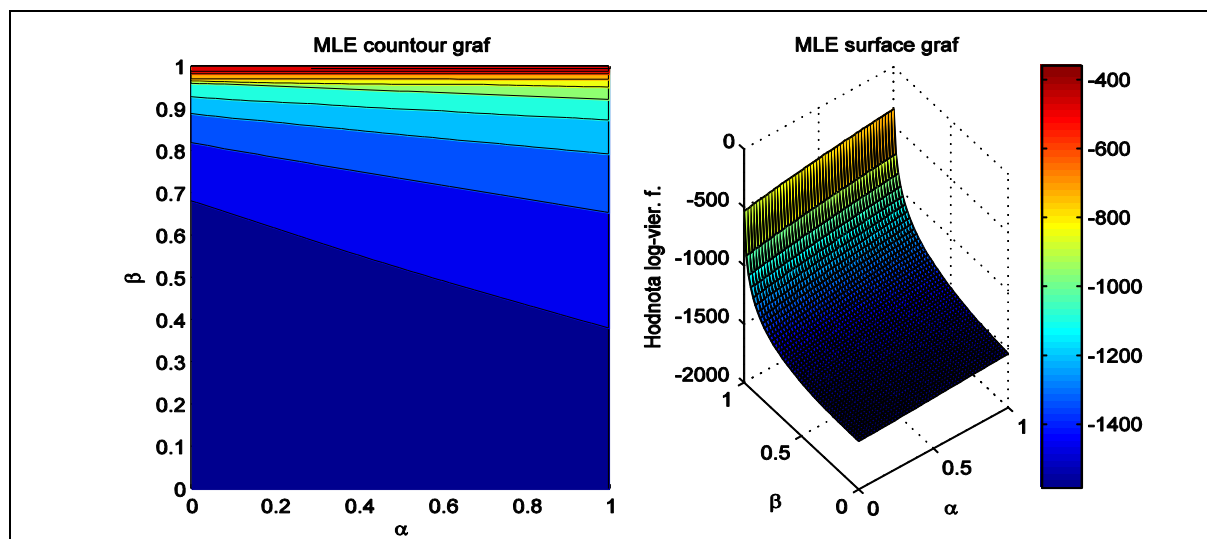
Obrázok 2 zobrazuje individuálne profily funkcie maximálnej vierohodnosti⁴⁵. Tvar funkcie maximálnej vierohodnosti je charakteristický pre danú kombináciu parametrov a s rastúcim počtom pozorovaní sa zásadne nemení. Vzhľadom na to, že do funkcie maximálnej vierohodnosti vstupujú jednotlivé pozorovania, tak je prirodzené, že s rastúcim počtom pozorovaní sa mení veľkosť a rozsah hodnôt, ktoré nado-
búda. Túto zmenu môžeme dobre vidieť na obrázku 3, ktorý zobrazuje hodnoty funkcie vierohodnosti

⁴ Z anglického „individual Profile Log-Likelihood Function“ Ma et al. (2007)

⁵ Odhad parametrov pomocou metódy maximálnej vierohodnosti bol realizovaný v prostredí Matlab, kde bola na hľadanie optima použitá funkcia fminsearch. Táto funkcia hľadá minimum funkcie a preto bola funkcia vierohodnosti transformovaná podľa pravidla $\max(l(\theta; w)) = \min(-l(\theta; w))$. Tento princíp je zachovaný aj pri zobrazovaní optima funkcie vierohodnosti v obrázku 2 a 3.

pre všetky možné kombinácie parametrov ARCH a GARCH, kde simulovaný proces zodpovedá vybranej kombinácii parametrov. Zároveň je zobrazená zmena hodnôt vierohodnostnej funkcie z hľadiska zvyšujúceho sa počtu pozorovaní zahrnutých do odhadu parametrov. Jednotlivé grafy nám umožňujú vzájomné porovnanie vzťahu medzi hodnotou vierohodnostnej funkcie a teoretickým počtom kombinácií parametrov, ktoré by bolo potrebné preskúmať v prípade hľadania optima pomocou explicitnej enumerácie kombinácií parametrov, kde s klesajúcimi hodnotami parametrov ARCH a GARCH klesá aj počet teoretických kombinácií parametrov, ktoré môžu byť optimom. Takisto je možné potvrdiť podľa tohto obrázka predpoklad o tom, že tvar funkcie vierohodnosti pri danej kombinácii parametrov nepodlieha zmene vplyvom väčšieho počtu pozorovaní, pretože nedochádza k veľkým zmenám možného počtu explicitných kombinácií pre najmenších 10 % hodnôt funkcie vierohodnosti.





Poznámka: Ľavý stĺpec zobrazuje contour (vrstevnicové) grafy a pravý stĺpec surface (povrchové) grafy pre individuálne profilové zobrazenie funkcie maximálnej vierohodnosti. Pre contour graf predstavujú jednotlivé farebne odlišené plochy 10 % z hodnôt funkcie maximálnej vierohodnosti. Každá dvojica grafov zodpovedá skúmanej kombinácii parametrov ARCH a GARCH. V tomto prípade to boli (.01,.97), (.4,.5) a (.01,01). Počet MC simulácií a počet pozorovaní je pre každú kombináciu parametrov ARCH a GARCH rovný 500.

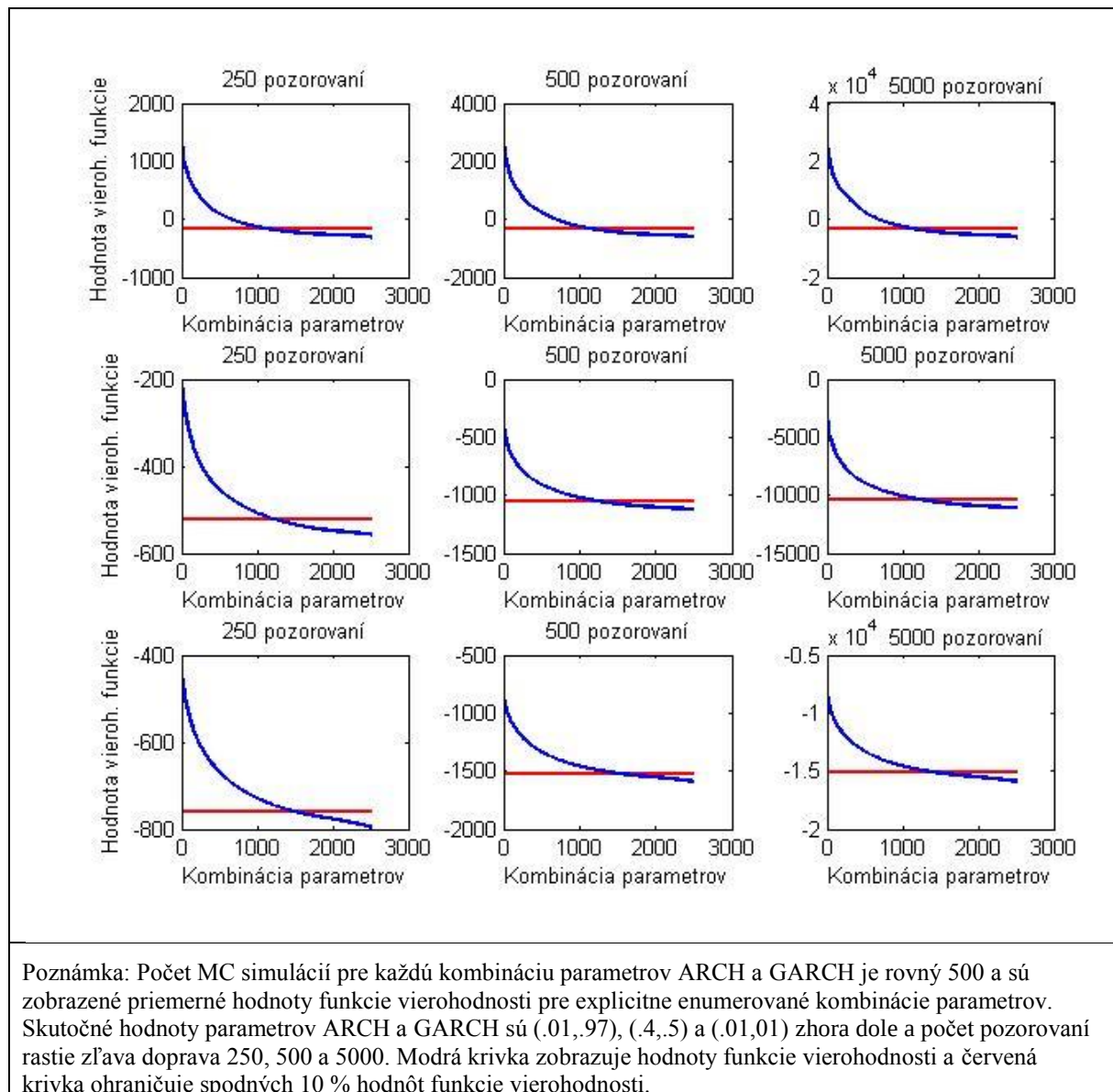
Obrázok 2: Tvar funkcie vierohodnosti pre vybrané kombinácie parametrov ARCH a GARCH

Záver

Hlavným cieľom článku bolo rozšírenie vedomostí o vlastnostiach odhadu modelu GARCH(1,1) s normálne rozdelenými náhodnými chybami odhadnutých metódou maximálnej vierohodnosti pri použití malého počtu pozorovaní. Inovatívne nadväzuje na prácu Hwang et al. (2006) a poskytuje komplexnejší obraz o skreslení odhadovaných parametrov vzhľadom na skutočné hodnoty modelu. Na základe výsledkov z Monte Carlo simulácií boli zostavené grafy zachytávajúce priemerné skreslenia jednotlivých parametrov modelu, ktoré sú spôsobené zahrnutím malého počtu pozorovaní do odhadu modelu. Charakter skreslenia je daný do širších súvislostí z hľadiska tvaru funkcie vierohodnosti v blízkosti optima, kombinácie parametrov modelu a počtu pozorovaní zahrnutých do odhadu modelu.

Dedikácia

Tento článok bol podporený grantom IGA F4/16 /2012, FIS, VŠE.



Obrázok 3: Vzťah medzi hodnotou funkcie vierohodnosti a počtom explicitne enumerovaných kombinácií parametrov

Literatúra

- Black, F. 1976. *Studies of stock market volatility changes*, Proceedings of the American Statistical Association, Business and Economic Statistics Section, 177–81.
- Bollerslev, T. 1986. *Generalized autoregressive conditional heteroscedasticity*. Journal of Econometrics **31**:307–327.
- Bollerslev, T., Wooldridge, J. M. 1992. *Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances*. Econometric Reviews **11**. 143-172.
- Chan, F., Theoharakis, B. 2009. *Re-Parameterization of Multi-regime STAR-GARCH Model*. MODSIM Congress, Cairns.
- Engle, R. 1982. *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation*. Econometrica **50**: 987–1008.
- Figlewski, S. 2004. *Forecasting volatility*. Financial Markets, Institutions and Instruments **6**, 1-88.

Fiorentini, G., Calzolari, G., Panattoni, L. 1996. *Analytic derivatives and the computation of GARCH estimates*. Journal of Applied Econometrics **11**: 399-417.

Hwang, S., Pereira, P. L. V. 2006. *Small sample properties of GARCH estimates and persistence*. The European Journal of Finance **12**, Issue 6-7.

Levy, G. 2001. *An introduction to GARCH models in Finance*. Financial Engineering News **22**.

Lumsdaine, R. L. 1995. *Finite-sample properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: a Monte Carlo investigation*. Journal of Business and Economic Statistics **13**, 1-10.

Ma, J., Nelson, Ch. R., Startz, R. 2007. *Spurious inference in the GARCH(1,1) model when it is weakly identified*. Studies in Nonlinear Dynamics and Econometrics **11**. 1-27.

JEL Classification: C58

Summary

GARCH(1,1), maximum likelihood method and sample size

This paper investigates GARCH (1,1) model parameters bias that is subject to small sample size. The model performance is investigated through various Monte Carlo simulations. Parameter estimates from maximum likelihood estimation method that are restricted under standard GARCH(1,1) restrictions are analyzed and the contour plot for mean bias parameter estimates is constructed. Such analysis is based on the assumption that the constant term parameter in unconditional variance formula does not influence ARCH and GARCH parameter in maximum likelihood estimation. We outline link between likelihood function shape, model parameters values and sample size included in estimation process. Several parameter combinations from different type of bias locations are scrutinized.

Key words: GARCH(1,1) model, maximum likelihood method, parameter bias, sample size.

Nekooperativní vícekritériální hry

Michaela Tichá

xticm11@vse.cz

Doktorandka oboru ekonometrie a operační výzkum

Školitel: doc. Mgr. Ing. Martin Dlouhý, Dr., MSc, (dlouhy@vse.cz)

Abstrakt: Práce se zabývá nekooperativní teorií vícekritériálních her. Jedná se o hry, kdy výplatní funkce alespoň jednoho z hráčů je vektor, a hráč chce maximalizovat všechna kritéria zároveň. Teorie hledání rovnovážných bodů se stále rozvíjí. Nalézt v obecné hře všechny rovnovážné body vícekritériální hry je dosud nevyřešený problém.

Nejprve práce představuje vícekritériální hry, uvádí základní definice. Dále zavádí dva typy rovnovážných bodů – silné a slabé – a naznačuje dosavadní možnosti řešení. Mezi ně patří skalarizace vektorové výplatní funkce, úplný popis řešení hry dvou hráčů s omezeným počtem strategií, nebo hledání tzv. ideálních rovnovážných bodů.

Nakonec je představen nový koncept řešení, který se zaměřuje na hledání rovnovážných bodů v čistých strategiích. Využívá prohledávání nejlepších odpovědí jednotlivých hráčů. Tímto způsobem lze velmi rychle a velmi snadno nalézt podmnožinu rovnovážných bodů v čistých strategiích. Výsledky jsou také snadno interpretovatelné, a tedy i snadno použitelné pro reálné případy. Vše je demonstrováno na konkrétním příkladu. Na závěr jsou shrnuty vlastnosti algoritmu.

Klíčová slova: vícekritériální hry, teorie her, vektorová výplatní funkce, nekooperativní hry

Úvod

Teorie her se zabývá řadou rozhodovacích situací, v nichž proti sobě stojí dva či více hráčů a každý se snaží maximalizovat svůj zisk. Kromě standardních konfliktů, kdy hráči maximalizují jedno kritérium, se často setkáváme se situacemi, kdy alespoň jeden z hráčů rozhoduje na základě více kritérií, které chce maximalizovat zároveň. Někdy lze zisk z kritérií jednoduše sečíst (např. prodává-li firma dva produkty a chce maximalizovat zisk z nich, lze vytvořit jediné kritérium jako součet zisků z jednotlivých produktů), ale často to nelze. Proto zde byla aplikována teorie vícekritériální optimalizace do teorie her a vznikla tak nová teorie vícekritériálních her. Ta se zabývá hledáním v určitém smyslu rovnovážných bodů her, kdy alespoň jeden hráč má více než jednu výplatní funkci. Z tohoto důvodu se vícekritériálním hrám také často říká hry s vektorovou výplatní funkcí.

Teorie vícekritériálních her je stále se rozvíjející disciplína. Způsob, kterým bychom v libovolné hře našli všechny rovnovážné body, zatím neexistuje. Je však řada konceptů, které naleznou buď nějakou podmnožinu rovnovážných bodů, nebo naleznou všechny rovnovážné body v nějaké specifické hře. V tomto článku představím vlastní způsob, jak snadno a rychle nalézt nějakou podmnožinu rovnovážných bodů.

Práce je převzata z (Tichá, 2012).

Základní koncepty

Nejprve zformulujeme nekooperativní vícekritériální hru více hráčů, poté zavedeme nerovnosti dvou vektorů a následně zadefinujeme rovnovážné body vícekritériální hry.

Definice 1. Nechť $n \in \mathbb{N}$ a $N = \{1, 2, \dots, n\}$ je množina n hráčů, X_i je množina čistých strategií hráče $i \in N$ a pro každého hráče $i \in N$ máme výplatní funkci

$$u_i : \prod_{j \in N} X_j \rightarrow \mathbb{R}^{r(i)},$$

která přiřazuje každé kombinaci strategií hráčů bod v $r(i)$ -dimenzionálním Euklidovském prostoru, kde $r(i)$ je počet kritérií hráče $i \in N$. Množinu

$$G = \langle N, (X_i)_{i \in N}, (u_i)_{i \in N} \rangle$$

nazveme *vícekriteriální hru*.

Obdobně provedeme rozšíření na smíšené strategie. Označme $\Delta(X_i)$ množinu smíšených strategií hráče $i \in N$. Pro smíšenou strategii $x_i \in \Delta(X_i)$, $j \in X_i$ označíme x_{ij} pravděpodobnost, že hráč i odpoví čistou strategií j .

Definice 2. Necht'

$$G = \langle N, (X_i)_{i \in N}, (u_i)_{i \in N} \rangle$$

je *vícekriteriální hra* z definice 1. Nyní pod množinou strategií $(X_i)_{i \in N}$ uvažujeme množiny smíšených strategií $\Delta(X_i)$ a hru G nazveme *smíšené rozšíření vícekriteriální hry*.

Dále budeme uvažovat pod pojmem *vícekriteriální hra* její smíšené rozšíření.

V další části budeme potřebovat porovnávat vektory, zavedeme tedy vektorové nerovnosti:

Definice 3. Necht' $a, b \in R^m$. Pak

$$a > b \quad \text{pokud } a_j \geq b_j \text{ pro všechna } j \in \{1, \dots, m\},$$

$$a \geq b \quad \text{pokud } a > b \text{ a } a \neq b,$$

$$a > b \quad \text{pokud } a_j > b_j \text{ pro všechna } j \in \{1, \dots, m\}.$$

Řekneme, že a *dominuje* b , pokud $a > b$.

Konečně se dostáváme k definici rovnovážného bodu *vícekriteriální hry*. Budeme používat obvyklé označení x_{-i} pro množinu strategií všech hráčů kromě hráče i : $x_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$.

V *jednokriteriální hře* pro Nashův rovnovážný bod platí, že je to strategický profil $x \in \prod_{i \in N} \Delta(X_i)$, pokud

žádný hráč nemá alternativní strategii $y_j \in \Delta(X_j)$, pro kterou by jeho výplatní funkce byla (za předpokladu stejných strategií protihráčů) vyšší, tj. $u_j(y_j, x_{-j}) > u_j(x)$. Ve *vícekriteriální hře* platí stejná definice. Výplatní funkce jsou vektorové, ale už máme zdefinovány vektorové nerovnosti, a tak můžeme přejít přímo k definici. Standardně se používají dvě definice – silný rovnovážný bod a slabý rovnovážný bod.

Definice 4. *Slabým rovnovážným bodem* nazveme strategický profil

$$x \in \prod_{i \in N} \Delta(X_i),$$

pokud pro žádného hráče $i \in N$ neexistuje strategie $\tilde{x}_i \in \Delta(X_i)$ taková, že

$$u_i(\tilde{x}_i, x_{-i}) > u_i(x_i, x_{-i}).$$

Definice 5. *Silným rovnovážným bodem* nazveme strategický profil

$$x \in \prod_{i \in N} \Delta(X_i),$$

pokud pro žádného hráče $i \in N$ neexistuje strategie $\tilde{x}_i \in \Delta(X_i)$ taková, že

$$u_i(\tilde{x}_i, x_{-i}) \geq u_i(x_i, x_{-i}).$$

Silný rovnovážný bod je v jistém smyslu analogie eficientního bodu ve *vícekriteriální optimalizaci*. Pro hráče je mnohem lepším rovnovážným bodem, protože říká, že pro žádného hráče neexistuje jiná strategie, při které by měl za daných strategických profilů protihráčů alespoň jednu výplatní funkci vyšší a ostatní stejné. Přirozeně, i když je jediné kritérium lepší, je pro hráče vektor výplatních funkcí lepší. Hledání silného rovnovážného bodu bývá ovšem mnohem náročnější, a tak se využívá definice slabého

rovnovážného bodu. Slabý rovnovážný bod nám říká pouze to, že pro žádného hráče neexistuje alternativní strategie, při které by měl za daných strategických profilů protihráčů všechny výplatní funkce vyšší.

Dále zavedeme definici nejlepší odpovědi, která odpovídá definici rovnovážného bodu.

Definice 6. Necht' $x_{-i} \in \Delta(X_{-i})$ je strategický profil všech hráčů kromě hráče i . Strategie $x_i \in \Delta(X_i)$ i -tého hráče se nazývá *nejlepší odpověď* i -tého hráče proti x_{-i} , pokud neexistuje strategie $\tilde{x}_i \in \Delta(X_i)$ taková, že platí $u_i(\tilde{x}_i, x_{-i}) > u_i(x_i, x_{-i})$. Speciálně v maticové hře dvou hráčů s maticemi M, N taková, že platí $\tilde{x}_1 M x_2 > x_1 M x_2$ pro prvního hráče, resp. $x_1 N \tilde{x}_2 > x_1 N x_2$ pro druhého hráče.

Potom platí, že rovnovážný bod je takový, že strategie jednotlivých hráčů jsou nejlepší odpovědi na strategické profily ostatních hráčů. V tomto znění definice 6 by se jednalo o slabé rovnovážné body, pokud bychom použili nerovnost $\geq$, jednalo by se o silné rovnovážné body.

Poznamenejme, že rovnovážných bodů může být a také často bývá nekonečně mnoho.

Nalézt všechny rovnovážné body je dosud nevyřešený problém. I kdybychom je uměli nalézt, při větší hře je obtížné popsat množinu rovnovážných bodů a následná interpretace je komplikovaná. Existují však některé koncepty možných řešení.

Skalarizace výplatních funkcí

Nejjednodušší a také nejlépe interpretovatelný způsob řešení je skalarizace vektorových výplatních funkcí. Pro každého hráče vytvoříme vektor vah jednotlivých kritérií a uděláme vážený součet, čímž vznikne jediné kritérium. Může se stát, že kritéria nejsou vzájemně srovnatelná, v takovém případě je potřeba je nejprve znormovat.

Následně známými způsoby vyřešíme standardní hru a potom platí, že nalezené řešení je zároveň slabý rovnovážný bod původní vícekritériální hry. Zároveň platí, že pro každý slabý rovnovážný bod existují vektory vah jednotlivých hráčů tak, že daný slabý rovnovážný bod je řešením vytvořené jednokritériální hry.

Toto řešení je popsáno v (Krus, Bronisz, 1994).

Struktura řešení u omezené hry

Při hře dvou hráčů takové, že alespoň jeden hráč má maximálně dvě strategie, existuje způsob, jak popsat celou množinu rovnovážných bodů. Vychází z toho, že nalezneme nejprve rozložení množiny na tzv. regiony stability, poté totéž pro druhého hráče a z nich pak vytváříme jednotlivé rovnovážné body. Způsob je popsán v (Borm, Vermeulen, Voorneveld, 2003).

Ideální rovnovážné body

Další varianta je definice ideálních rovnovážných bodů.

Definice 7. *Ideálním rovnovážným bodem* nazveme strategický profil

$$x \in \prod_{i \in N} \Delta(X_i),$$

pokud pro žádného hráče $i \in N$ neexistuje strategie $\tilde{x}_i \in \Delta(X_i)$ taková, že

$$u_i(\tilde{x}_i, x_{-i}) > u_i(x_i, x_{-i}).$$

Jedná se tedy o takové body, kdy výplatní funkce je nejlepší ve všech kritériích zároveň. Takové body lze nalézt všechny. Hledáme jako průnik několika standardních Nashových rovnovážných bodů zvlášť přes tzv. reprezentativní množinu vah vektorů. Pokud existují, pak tak jsou zároveň slabými i silnými rovnovážnými body. Bohužel však mnohem častěji neexistují a jedná se spíše o speciální případ. Způsob hledání je popsán v (Voorneveld, Grahn, Dufwenberg, 2000).

Řešení v čistých strategiích

Ve standardní hře je řešení v čistých strategiích spíše výjimečné a existuje pouze ve speciálních případech. Ve hře s vektorovou výplatní funkcí je však situace jiná. Řešení v čistých strategiích naopak velice často existuje. A jeho existence je tím pravděpodobnější, čím méně čistých strategií a čím více kritérií hráči mají. Je to způsobeno velkým množstvím rovnovážných bodů, což vyplývá z toho, že eficientních bodů u vícekritériální optimalizace také bývá velmi mnoho.

Nalezení několika řešení v čistých strategiích je velice rychlé a snadné. Vyjdeme z toho, že rovnovážný bod je takový, že pro každého hráče je nejlepší odpovědí na strategický profil ostatních. Předpokládáme-li pouze řešení v čistých strategiích, lze pro každého hráče nalézt nejlepší odpovědi na všechny možné kombinace čistých strategií ostatních hráčů z hlediska jednotlivých kritérií, kterých je konečný počet. Pokud je nejlepší odpovědí na nějaký strategický profil protihráčů čistá strategie X z hlediska jednoho kritéria a čistá strategie Y z hlediska druhého kritéria, pak jsou nejlepší odpovědi na daný strategický profil protihráčů obě čisté strategie. To vychází z definice rovnovážných bodů. Ukážeme si to na příkladu dvou hráčů.

Příklad 1. Uvažujme maticově zadanou hru dvou hráčů. První hráč má tři kritéria, druhý hráč má dvě kritéria. Hráč 1 má strategie X a Y , hráč 2 má strategie A , B , C . Hodnoty výplatních funkcí jsou zobrazeny v následujících maticích.

První hráč:

$$\begin{bmatrix} (2,1,2) & (4,2,9) & (0,2,1) \\ (3,7,1) & (7,1,0) & (1,8,2) \end{bmatrix}$$

Druhý hráč

$$\begin{bmatrix} (0,4) & (2,5) & (1,2) \\ (1,7) & (7,3) & (1,0) \end{bmatrix}$$

Nyní budeme zkoumat nejlepší odpovědi hráčů. Nejprve se podíváme na nejlepší odpovědi prvního hráče. Rozepíšeme si matici vektorů výplatních funkcí na jednotlivé matice a nahlédneme, jaké čisté strategie jsou nejlepší odpovědi na jednotlivé čisté strategie druhého hráče.

První kritérium:

$$\begin{bmatrix} 2 & 4 & 0 \\ 3 & 7 & 1 \end{bmatrix}.$$

Druhé kritérium:

$$\begin{bmatrix} 1 & 2 & 2 \\ 7 & 1 & 8 \end{bmatrix}.$$

Třetí kritérium:

$$\begin{bmatrix} 2 & 9 & 1 \\ 1 & 0 & 2 \end{bmatrix}.$$

Pokud druhý hráč zahraje strategii A , nejlepší odpovědi prvního hráče je jak strategie X , tak strategie Y . Vycházíme z úvahy, že taková čistá strategie, která je alespoň v jednom kritériu maximální (vzhledem k strategii A druhého hráče), je nejlepší odpovědí. Je tomu tak proto, že taková strategie nemůže být dominována žádnou jinou, v daném kritériu bude mít vždy vyšší hodnotu než jakákoli kombinace ostatních strategií.

Tedy za předpokladu strategie A druhého hráče je dle prvního kritéria nejlepší odpovědí strategie Y , dle druhého kritéria strategie Y a dle třetího kritéria strategie X . Takže na strategii A je nejlepší odpovědí jak strategie X , tak strategie Y .

Obdobně zjistíme, že za předpokladu strategie B druhého hráče je nejlepší odpovědí strategie X i Y . Za předpokladu strategie C je nejlepší odpovědí pouze strategie Y , neboť dominuje strategii X ve všech kritériích.

Nyní provedeme analogickou analýzu ze strany druhého hráče. Nejprve si rozepíšeme matici vektorů na matice dle jednotlivých kritérií:

První kritérium:

$$\begin{bmatrix} 0 & 2 & 1 \\ 1 & 7 & 1 \end{bmatrix}.$$

Druhé kritérium:

$$\begin{bmatrix} 4 & 5 & 2 \\ 7 & 3 & 0 \end{bmatrix}.$$

Pokud první hráč zahraje strategii X , pak nejlepší odpovědí druhého hráče je pouze čistá strategie B , neboť v obou kritériích dominuje strategii A i C . Nejlepší odpovědí na strategii Y jsou pak čisté strategie A a B .

Protože platí, že dvojice strategií (x_1, x_2) je rovnovážným bodem právě tehdy, když x_1 je nejlepší odpovědí na x_2 a x_2 je nejlepší odpovědí na x_1 , snadno z předchozích úvah určíme některé rovnovážné body.

Rovnovážnými body v čistých strategiích tak jsou dvojice strategií (X, B) , (Y, A) , (Y, B) . To plyne z toho, že jsme zjistili, že X je nejlepší odpovědí na B i B je nejlepší odpovědí na X , obdobně i pro zbylé dva body. Prakticky okamžitě jsme tak našli několik rovnovážných bodů.

⊕

Zhodnocení

Hlavní výhodou výše popsaného způsobu je jeho nenáročnost. Jedná se o snadný algoritmus, pomocí něhož velice rychle nalezneme několik rovnovážných bodů. Jen tak mimochodem jsou to řešení v čistých strategiích. Příklad byl vytvořen pro dva hráče, platí však pro obecný počet hráčů n .

Zároveň lze hledat jak silné, tak slabé rovnovážné body. Jediný rozdíl je v určování nejlepších odpovědí. Pro jednoduchost si to ukážeme na dalším příkladu.

Příklad 2. V příkladu 1 nedošlo nikde ke shodě hodnot u více strategií. Uvažovali jsme dominanci ve smyslu „ $\leq$ “ z definice 3 a všechny nalezené rovnovážné body byly silné. Pokud však k nějaké shodě dojde, můžeme se držet definice dominance a hledat tak silné rovnovážné body, nebo můžeme chápat dominanci ve smyslu „ $\leq$ “ a hledat slabé rovnovážné body. Vráťme se k předchozímu příkladu a malinko ho pozměníme.

První hráč:

$$\begin{bmatrix} (2,1,2) & (7,2,9) & (0,2,1) \\ (3,7,1) & (7,1,9) & (1,8,2) \end{bmatrix}.$$

Druhý hráč:

$$\begin{bmatrix} (0,4) & (2,5) & (1,2) \\ (1,7) & (7,3) & (1,0) \end{bmatrix}.$$

Změnili jsme hodnoty výplatních funkcí prvního hráče u prvního a třetího kritéria při strategii B druhého hráče. Opět budeme hledat nejlepší odpovědi. Vzhledem k nezměněné matici druhého hráče zůstávají i jeho nejlepší odpovědi stejné.

U prvního hráče zůstávají stejné nejlepší odpovědi na strategie A a C .

Situace je zajímavější za předpokladu volby strategie B druhým hráčem. U prvního a třetího kritéria jsou hodnoty stejné u strategie X i Y . U druhého kritéria X dominuje Y . Ted' záleží, zda hledáme slabé či silné rovnovážné body.

Pokud nás zajímají slabé rovnovážné body, pak při rovnosti více strategií považujeme za nejlepší odpovědi obě. Pokud nás naopak zajímají silné rovnovážné body, pak při rovnosti více strategií nepovažujeme (u daného kritéria) za nejlepší odpověď ani jednu, ovšem s jednou výjimkou. Pakliže by byla rovnost u všech kritérií, pak také považujeme všechny strategie za nejlepší odpovědi i za předpokladu hledání silných rovnovážných bodů. Vše snadno vychází z definice slabého a silného rovnovážného bodu.

Tedy v našem příkladě uvažujeme-li slabé rovnovážné body, zůstanou jimi dvojice strategií (X, B) , (Y, A) , (Y, B) . Uvažujeme-li silné rovnovážné body, situace se změní. Při volbě strategie B druhým hráčem je strategie Y dominována strategií X , nejlepší odpovědi na strategii B je tak pouze strategie Y . Tím se nám ztratí rovnovážný bod (X, B) a zůstanou pouze rovnovážné body (Y, A) , (Y, B) .

⊕

Nakonec ještě poznamenejme, že existence není zaručena u libovolné vícekritériální hry. Zároveň tento způsob nenalezne nutně všechna řešení v čistých strategiích. Protipříklady lze nalézt v (Tichá, 2012).

Závěr

V článku byl předveden nový způsob hledání rovnovážných bodů vícekritériální hry, a to speciálně v čistých strategiích. Jedná se o nenáročný algoritmus, který může být snadno převeden do nějakého programovacího jazyka a využit tak v hledání řešení libovolně velkých her. Řešení v čistých strategiích je také nekomplikované a jednoduše interpretovatelné. Výsledky tak mohou být snadno převedeny do praxe, na rozdíl od složitého úplného popisu všech rovnovážných bodů. Jelikož je v (Tichá, 2012) předvedeno, že způsob nemusí nutně nalézt všechna taková řešení, je předmětem dalšího zkoumání, jak nalézt všechna řešení v čistých strategiích. K tomu by stačilo nalézt podmínku, kdy určitý rovnovážný bod tímto způsobem nenalezneme.

Literatura

BORM, Peter, VERMEULEN, Dries, VOORNEVELD, Mark, 2003. The Structure of the Set of Equilibria for Two Person Multicriteria Games. *European Journal of Operation Research*, Vol. 148, pp. 480-493.

KRUS, Lech, BRONISZ, Piotr, 1994. On n -person Noncooperative Multicriteria Games Described in Strategic Form. *Annals of Operation Research*, Vol. 51, pp. 83-97

TICHÁ, Michaela, 2012. Vícekritériální hry. Diplomová práce, MFF UK.

VOORNEVELD, Mark, GRAHN, Sofia, DUFWENBERG, Martin, 2000: Ideal Equilibria in Noncooperative Multicriteria Games, *Mathematical Methods of Operations Research*, Vol. 52, pp. 65-77.

JEL Classification: C72

Dedikace: Příspěvek je zpracován jako součást výzkumného projektu interní grantové agentury F4/19/2013.

Summary

Noncooperative multicriteria games

The concern of the thesis is to discuss noncooperative multicriteria games. Multicriteria game is the special case from the game theory if the payoff function of at least one player is a vector and the player wants to maximize all the criteria at the same time. The theory of searching equilibria is still in development. To find all equilibria in a general multicriteria game is an unsolved problem.

The multicriteria games and basic definition are introduced in the beginning. Subsequently there are shown two types of equilibria – strong and weak – and possibilities of their solution. It is for example scalarization of the vector payoff, full description of the game of two players with limited number of strategies and the concept of ideal equilibria.

Finally a new concept of solution is introduced. It focus on looking for equilibria in pure strategies. Searching according to best replies of single player is used. It is possible to find the subset of equilibria in pure strategies pretty fast and easily. The results are also clear for interpretation and they can be used for real cases. Everything is demonstrated on numerical example. At the end the properties of the algorithm are summarized.

Key words: multicriteria games, game theory, vector payoffs, noncooperative games.

Demografický a ekonomický pohľad na populačné starnutie

Kornélia Cséfalvaiová

xcsek00@vse.cz

Doktorandka oboru statistika

Školiteľ: doc. Ing. Jitka Langhamrová, CSc., (langhamj@vse.cz)

Abstrakt: Proces populačného starnutia je v súčasnosti veľmi rozoberanou problematikou a vyvoláva mnoho diskusií. Proces demografického starnutia súvisí so zmenami vo vekovej štruktúre obyvateľstva, pričom najintenzívnejší nárast sledujeme pri podiele starších osôb na úkor detskej zložky populácie. Ustavične sa predlžuje stredná dĺžka života, zlepšuje sa úmrtnosť predovšetkým v staršom veku a plodnosť je na hodnotách, ktoré nezaručujú prostú reprodukciu obyvateľstva. V dnešnej vyspelej spoločnosti, vďaka pokrokom v oblasti medicíny a zdravotnej starostlivosti, nebude viditeľná iba zmena nárastu zložky 60–70 ročných osôb, ale postupne sa bude zvyšovať aj podiel 80–90 ročných, či dokonca 100 ročných a starších osôb. Meniaca sa veková skladba populácie Českej republiky bude pre ekonomiku 21. storočia predstavovať mnohé výzvy. V budúcnosti bude za potreby venovať nemalú časť štátnych príjmov sociálnym a zdravotným službám. Taktiež sa musí radikálne zmeniť charakter dôchodkového systému. Práca prezentuje vývoj vybraných demografických ukazovateľov súvisiacich s populačným starnutím v krajinách Vyšehradskej štvorky a ekonomické problémy spojené so starnúcou spoločnosťou.

Kľúčová slova: Populačné starnutie, Veková štruktúra obyvateľstva, Index závislosti starých, Stredná dĺžka života.

Úvod

V posledných desaťročiach sa podiel produktívnej zložky obyvateľstva a ekonomicky závislých osôb (osoby v predproduktívnom veku a osoby v poproduktívnom veku) výrazne nezmenil, avšak pri druhej spomínanej skupine došlo k výraznému preskupeniu – podiel staršieho obyvateľstva vzrástol v takej miere ako poklesol podiel detskej zložky. Kritický okamžik nastane vtedy, keď poklesne podiel osôb v produktívnom veku, ktorí sú ekonomicky aktívni. Zároveň tento znížený podiel ekonomicky aktívneho obyvateľstva bude musieť užiť rastúci počet ekonomicky závislých osôb, ktorých prevažnú časť budú tvoriť práve osoby v dôchodkovom veku.

Pre hodnotenie tohto procesu sa veľmi často používajú ukazovatele ako index hospodárskeho zaťaženia a indexy závislosti mladých či starých. Index závislosti používame, keď chceme zistiť pomer osôb v produktívnom veku a osôb v predproduktívnom a poproduktívnom veku, teda koľko závislých osôb musí užiť jedna produktívna osoba (Langhamrová, 2007).

Často sa uvádza, že dôsledky populačného starnutia a zastavenie demografického starnutia môžu byť vykompenzované vyšším počtom narodených detí. Demografické charakteristiky však nebudú v rovnováhe iba samotným zvýšením plodnosti a pôrodnosti. Všetky krajiny by mali podporovať rodiny s deťmi, ale taktiež je potreba sa zamyslieť nad financovaním dôchodkového systému, pretože chýbajúci rozdiel v penzijnom fonde musí byť vyrovnaný čoraz vyššími čiastkami zo štátneho rozpočtu a táto suma bude rozhodne chýbať napríklad v oblasti ekonomického rozvoja. Narastajúci podiel staršieho obyvateľstva bude vyžadovať pripravenosť sociálneho a zdravotného systému, rozšírenie nových inštitúcií, domovov dôchodcov a sociálnych služieb.

Dlhodobým riešením by mohol byť pružnejší prechod z obdobia ekonomickej aktivity do obdobia ekonomickej nečinnosti, teda možnosť seniorov pracovať na čiastočný úväzok alebo možnosť využitia iných skrátených pracovných úväzkov. Zároveň vytvorenie vhodných pracovných podmienok môže prispieť k dožitiu vyššieho veku v dobrom zdraví.

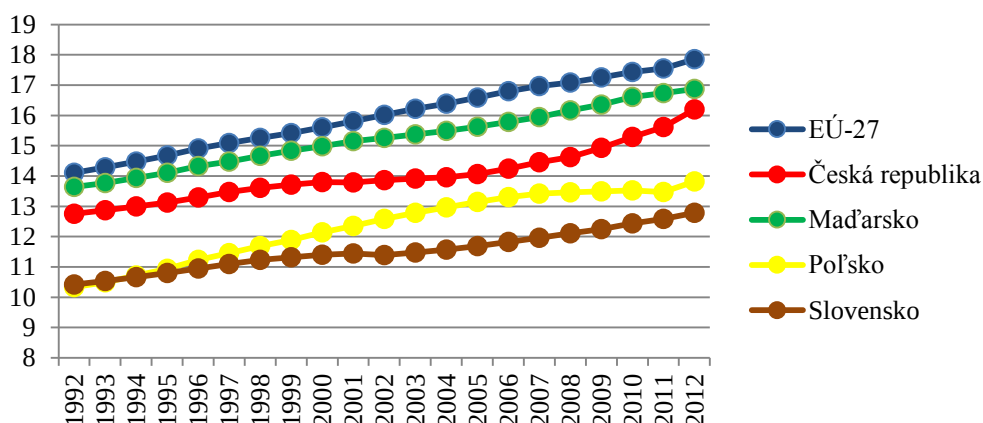
Na vyváženie populačného vývoja je potrebný paralelný vývoj viacerých procesov. Krajiny bývalého socialistického bloku z pohľadu demografického rozvoja mierne zaostávajú za krajinami vyspelejšími. Prejavuje sa to predovšetkým v kratšej dĺžke života ako mužov, tak žien. Pre zmiernenie demografického starnutia by bol potrebný zvýšený počet novorodencov. V súčasnej dobe sa však nepredpokladá, že by sa počet narodených v ČR zvyšoval, skôr bude dochádzať k poklesu vďaka tomu, že budú rodiť

slabšie ročníky žien v plodivom veku a plodnosť sa výrazne nezvyší. Môžeme predpokladať, že sa i naďalej bude zvyšovať stredná dĺžka života mužov i žien v Českej republike, Maďarsku, Poľsku a na Slovensku. Aj keď bude dochádzať k úbytku obyvateľstva prirodzeným prírastkom, nemusí dochádzať k úbytku obyvateľstva a to vtedy, keď bude kladné migračné saldo.

Vybrané demografické a ekonomické aspekty populačného starnutia v krajinách Vyšehradskej štvorky

Populačné starnutie je výsledkom demografických procesov, pri ktorých dochádza jednak k predlžovaniu strednej dĺžky života a na druhej strane klesá úroveň plodnosti, čoho dôsledkom je, že podiel starých osôb prevyšuje podiel detí.

Pri porovnaní podielu osôb 65 ročných a starších na celkovom obyvateľstve v krajinách Vyšehradskej štvorky s priemerom členských štátov Európskej únie (ďalej v texte budeme uvažovať EÚ–27), vidíme značné rozdiely. Za celú EÚ–27 tvoril podiel starých obyvateľov 17,9 % v roku 2012. Zo skúmaných krajín sa najväčší podiel seniorov vyskytuje v Maďarsku. Je to dôsledok najmä vplyvu plodnosti v minulosti, zlepšujúcich sa úmrtnostných pomerov a podobne (viď obrázok 1).

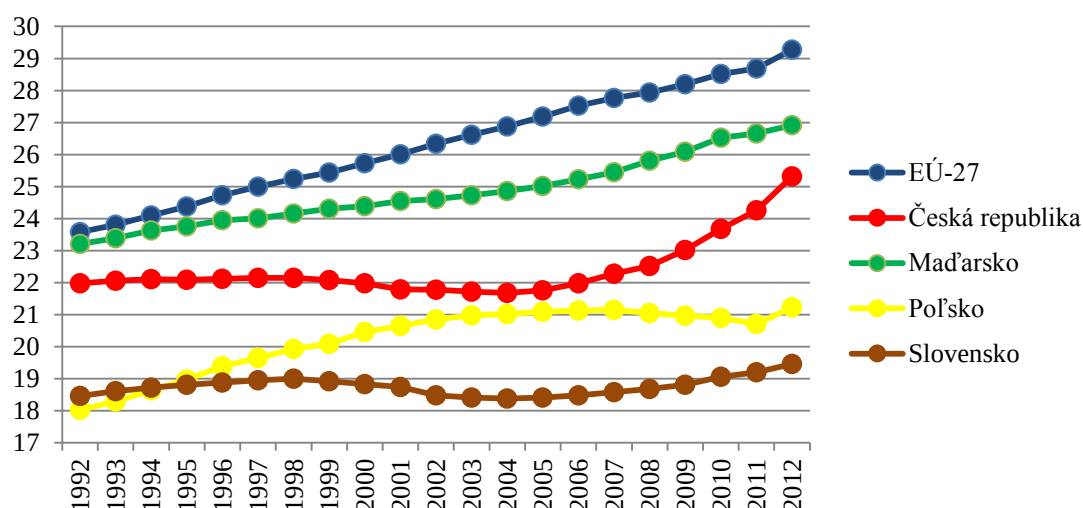


Obrázok 26: Podiel osôb 65 ročných a starších v krajinách Vyšehradskej štvorky a Európskej únie v období 1992–2012 (%)

Zdroj: Eurostat, vlastná konštrukcia

Okrem podielu osôb 65-ročných a starších sa na hodnotenie stavu populačného starnutia používajú ukazovatele index závislosti starých a index staroby. Index závislosti starých predstavuje podiel poproduktívnej zložky populácie (65+) a produktívnej zložky populácie (20–64 rokov). Index závislosti starých spolu s indexom závislosti mladých tvoria index závislosti, ktorý vyjadruje, koľko predproduktívnych (0–19 ročných) a poproduktívnych (65+) osôb pripadá na osobu v produktívnom veku (20–64 rokov). Kým index závislosti mladých sa za posledných 20 rokov v krajinách EÚ znížil, opačný vývoj nastal u indexu závislosti starých. Index staroby vyjadruje pomer prarodičov a detí (0–14 ročných) v populácii, teda koľko starých osôb z prarodičovskej generácie pripadá na jedno dieťa, respektíve na 100 detí. Vývoj indexu závislosti starých a indexu staroby za posledných dvadsať rokov v Českej republike, Maďarsku, Poľsku a na Slovensku je znázornený na obrázku 2 a 3.

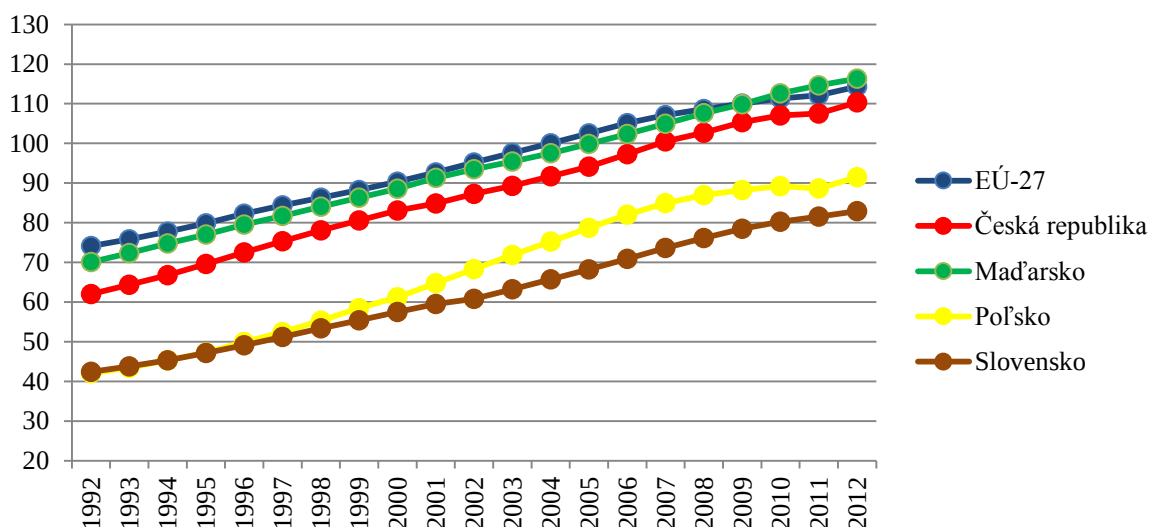
Z obrázku 2 vyplýva, že úroveň indexu závislosti starých i v tomto prípade zaostáva za priemerom Európskej únie. Index závislosti starých sa v EÚ–27 zvýšil z 23,6 % v roku 1992 na 29,3 % v roku 2012. Z porovnávaných krajín bol najnižší pomer starých osôb vzhľadom k osobám v produktívnom veku na Slovensku (19,5 % v roku 2012). Časový rad Slovenska mal veľmi stabilný priebeh za posledné dve desaťročia – medzi rokom 1992 a 2012 sa index závislosti starých zvýšil o jeden percentuálny bod. Naopak, v rovnakom časovom období sa hodnota indexu v Maďarsku zvýšila o 3,7 percentuálne body. Maďarsko sa výrazne líši od ostatných krajín Vyšehradskej štvorky, a tento rozdiel bol spôsobený jednak odlišným rokom transformácie systému a taktiež súčasnými rozdielmi, ktoré sú väčšie ako v roku 1992. V Českej republike sa index závislosti starých zvýšil z hodnoty 22,0 % v roku 1992 na 25,3 % v roku 2012 – v českej populácii, rovnako ako v prevažnej väčšine európskych krajín, stúpa podiel starších obyvateľov.



Obrázok 2: Index závislosti starých v krajinách Vyšehradskej štvorky a Európskej únie v období 1992–2012 (%)

Zdroj: Eurostat, vlastná konštrukcia

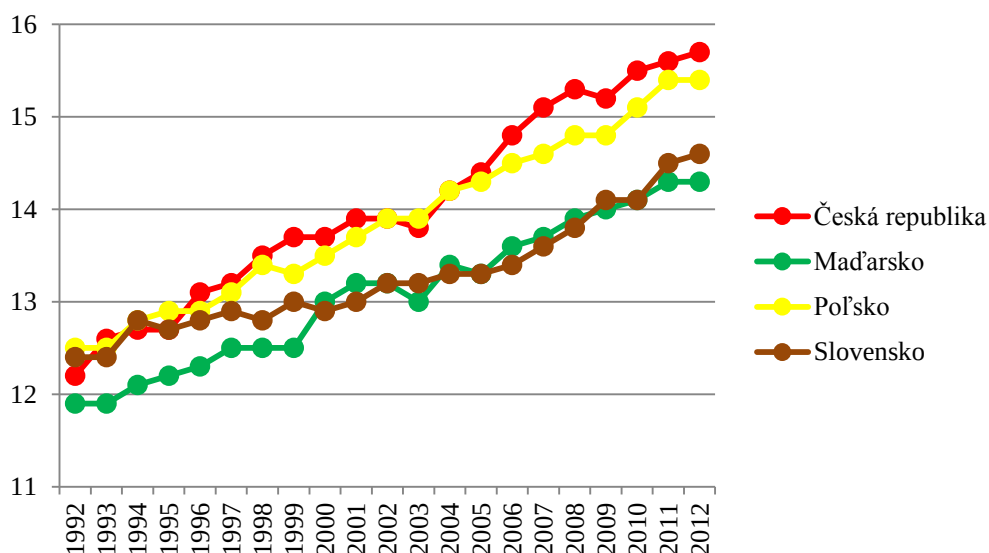
Podobný trend vývoja ako pri indexe závislosti starých osôb môžeme sledovať aj pri indexe staroby (viď obrázok 3). Radikálny pokles počtu narodených detí spôsobil nárast indexu staroby v sledovaných krajinách. Z výsledkov je zrejmé, že v Českej republike pripadá jedenásť starých osôb na desať detí (hodnota indexu staroby 110,4 v roku 2012). Vývoj hodnoty indexu staroby v Maďarsku je podobný ako v Českej republike a v EÚ (114,3 % v roku 2012). Na Slovensku pripadá osem prarodičov na desať detí a v Poľsku deväť starých osôb na desať detí.



Obrázok 3: Index staroby v krajinách Vyšehradskej štvorky a Európskej únie v období 1992–2012 (%)

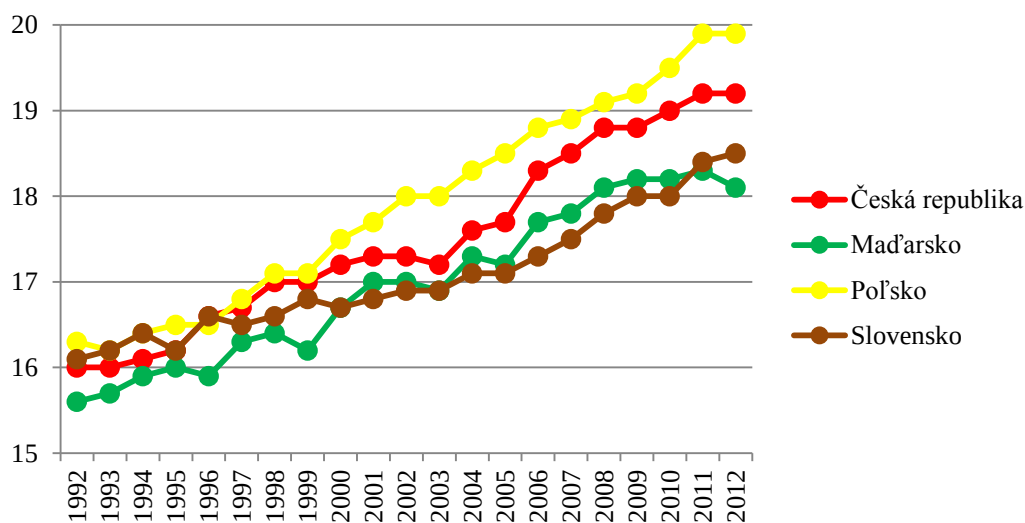
Zdroj: Eurostat, vlastná konštrukcia

Súčasne dochádza k zlepšovaniu zdravotného stavu obyvateľstva a k predlžovaniu strednej dĺžky života (nádeje dožitia). Vývoj hodnôt strednej dĺžky života je rozdielny pre mužov a ženy z dôvodu vyšších mier úmrtnosti mužov. Z obrázku 4 vidíme, že keby sa úmrtnostné pomery nemenili v priebehu ďalšieho života, 65-ročný muž by v priemere prežil ďalších 15,7 rokov v Českej republike. Vo všetkých sledovaných krajinách sledujeme rastúci trend strednej dĺžky života.



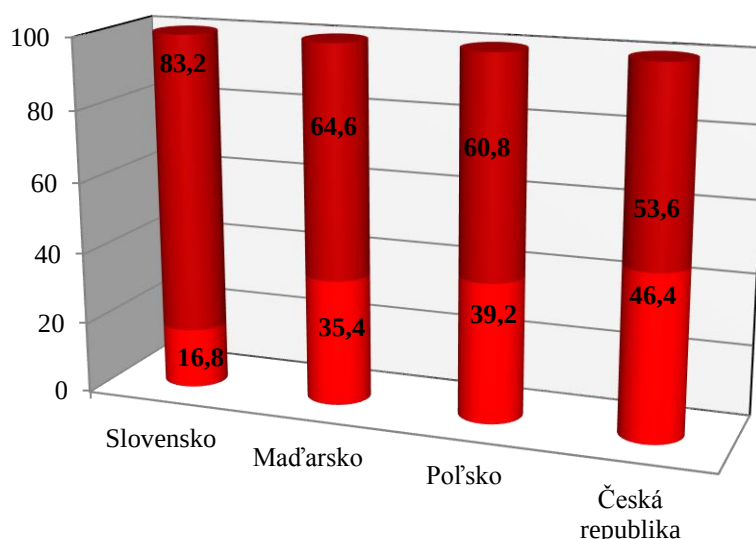
Obrázok 4: Vývoj strednej dĺžky života mužov vo veku 65 v krajinách Vyšehradskej štvorky v období 1992–2012 (%)
Zdroj: Eurostat, vlastná konštrukcia

Stredná dĺžka života pre 65-ročné ženy je vyššia ako pre mužov, najvyššia hodnota bola zaznamenaná v Poľsku (19,9 v roku 2012). Na poslednom mieste sa nachádza Maďarsko, a to rovnako ako v prípade mužov, tak i žien. Nepriaznivý vývoj strednej dĺžky života maďarskej populácie v rámci stredovýchodnej Európy existoval už aj pred spoločenskou transformáciou (Széman, 2013).



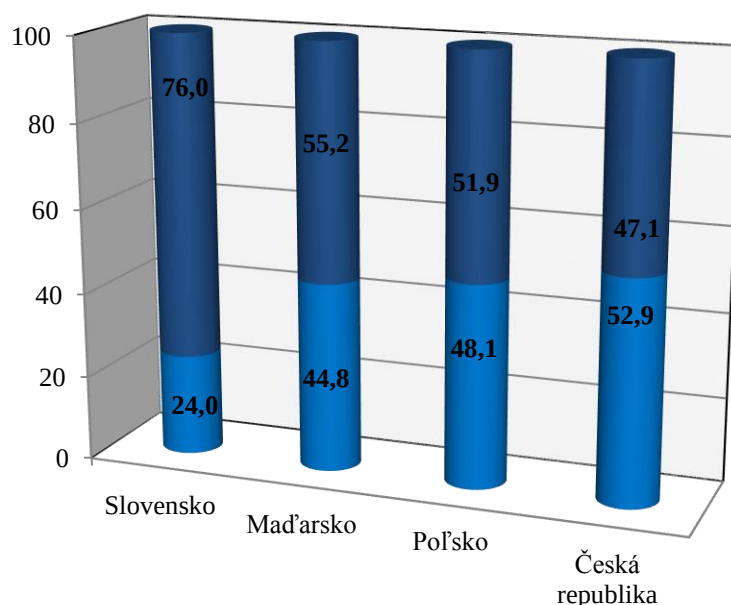
Obrázok 5: Vývoj strednej dĺžky života žien vo veku 65 v krajinách Vyšehradskej štvorky v období 1992–2012 (%)
Zdroj: Eurostat, vlastná konštrukcia

Z hľadiska predlžovania nádeje dožitia je dôležité, či pridané roky života strávi osoba v dobrom alebo zlom zdraví. Nádej dožitia podľa zdravotného stavu sa v tejto súvislosti sleduje pomocou ukazovateľa nádej dožitia v zdraví a vyjadruje počet rokov, ktoré v priemere ostávajú osobe v určitom veku k prežitiu v dobrom zdraví, teda bez zdravotného obmedzenia. Na obrázkoch 6 a 7 je znázornený podiel zdravých rokov života na strednej dĺžke života pre ženy a mužov vo veku 65 rokov. Hodnoty zdravej dĺžky života boli najnižšie na Slovensku (3,1 rokov pre ženy a 3,5 rokov pre mužov), naopak najvyššie v Českej republike (8,9 rokov pre ženy a 8,3 rokov pre mužov). Rovnaký záver zobrazuje i obrázok 6 a 7, na ktorých vidíme, že podiel zdravej dĺžky života v ČR v roku 2012 činil 46,4 % pre ženy a 52,9 % pre mužov. Z obrázkov 6 a 7 je viditeľný i percentuálny podiel rokov prežitých so zdravotným obmedzením, ktorý bol zo sledovaných krajín najvyšší na Slovensku (83,2 % pre 65-ročné ženy a 76 % pre 65-ročných mužov).



Obrázok 6: Podiel zdravej dĺžky života žien vo veku 65 rokov na strednej dĺžke života žien vo veku 65 rokov v roku 2012 (%)

Zdroj: Eurostat, vlastný výpočet

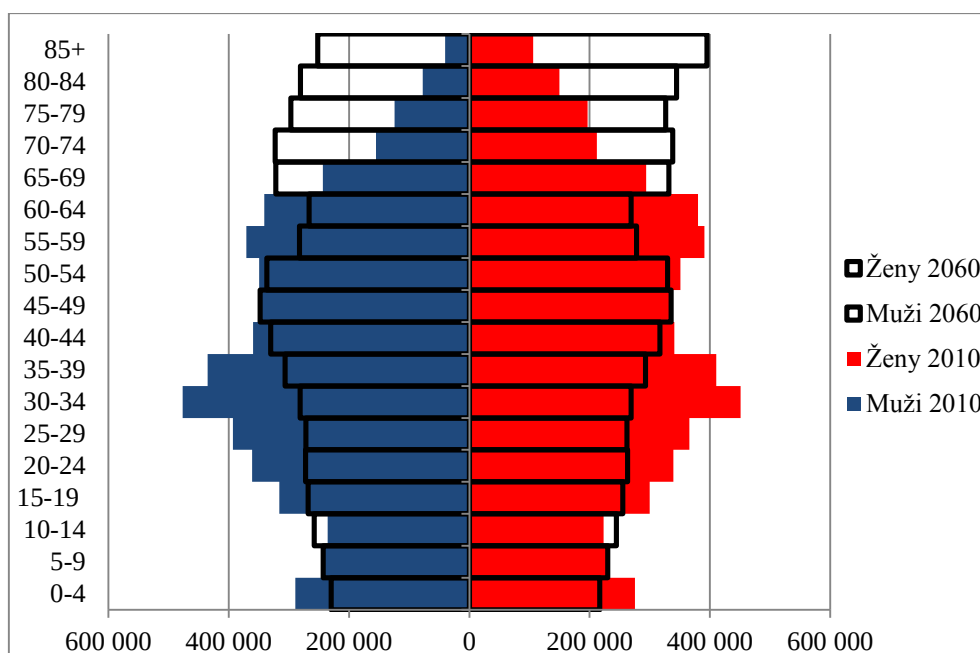


Obrázok 7: Podiel zdravej dĺžky života mužov vo veku 65 rokov na strednej dĺžke života mužov vo veku 65 rokov v roku 2012 (%)

Zdroj: Eurostat, vlastný výpočet

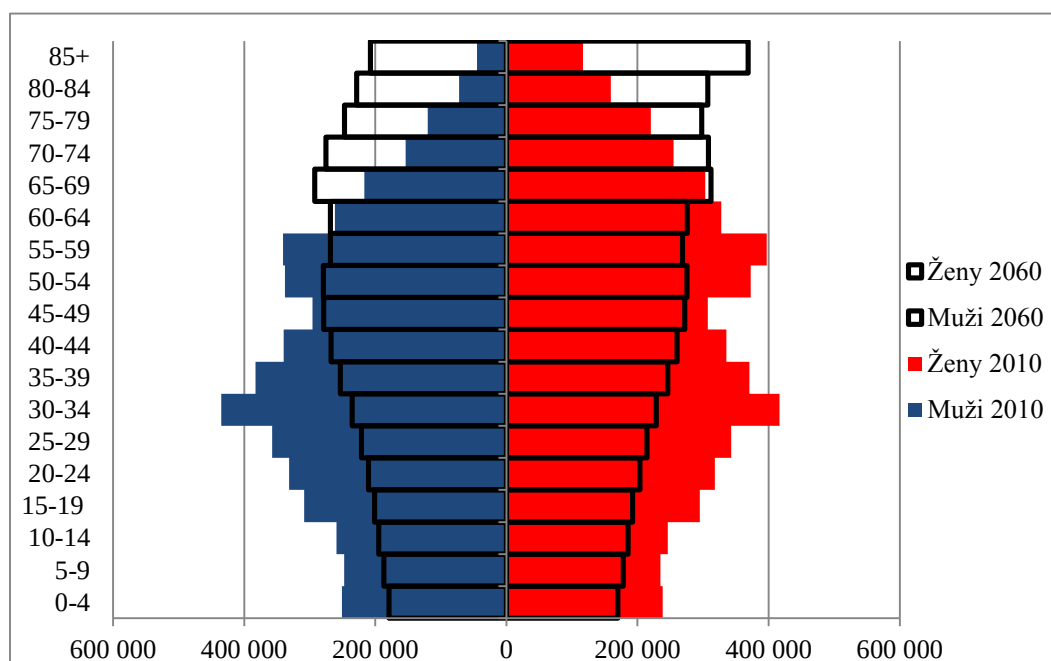
Tak ako sa mení veková štruktúra populácie, je dôležité sledovať, aká je veková a pohlavná štruktúra populácie vo veku ekonomickej aktivity, zaujíma nás potenciálna pracovná sila. Demografické charakteristiky sú z hľadiska pracovnej sily vstupným faktorom ovplyvňujúcim výslednú ekonomickú aktivitu. Súčasný stav v Českej republike, Poľsku a na Slovensku je charakterizovaný značným poklesom pôrodnosti a súbežne i prechodom početnej generácie narodennej v 70. rokoch 20. storočia do produktívneho veku, ktorý zapríčinil výrazné zvýšenie podielu vekovej skupiny 20–64. V Maďarsku tento proces začal skôr ako v ostatných troch krajinách a z tohto dôvodu je populačné starnutie v pokročilejšom štádiu ako v Českej republike, Poľsku a na Slovensku.

Súčasný stav, ktorý je charakterizovaný vysokým podielom ekonomicky aktívnych osôb a vysoko kvalifikovanou pracovnou silou môže predstavovať významný potenciál pre ekonomický rast a konkurencieschopnosť ekonomiky (Slaný, 2007). Tento stav, v ktorom sa pozitívne využijú zmeny vo vekovej štruktúre obyvateľstva, sa nazýva demografická dividenda (taktiež dar, okno).



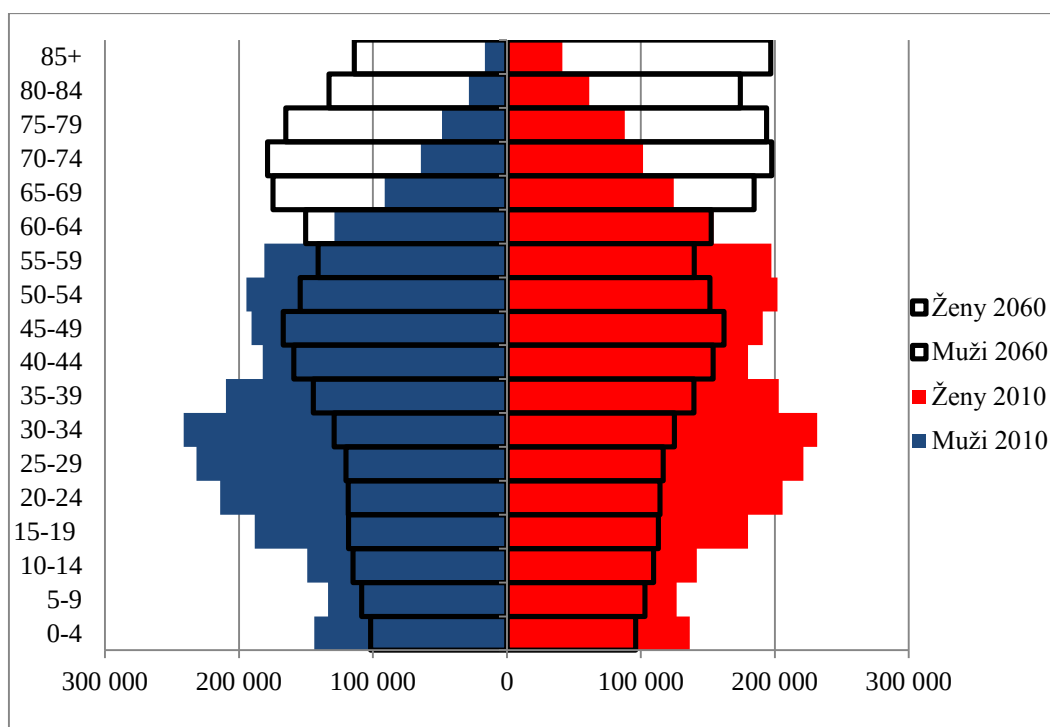
Obrázok 8: Porovnanie vekového zloženia obyvateľstva Českej republiky v roku 2010 a 2060

Zdroj: Eurostat, vlastná konštrukcia



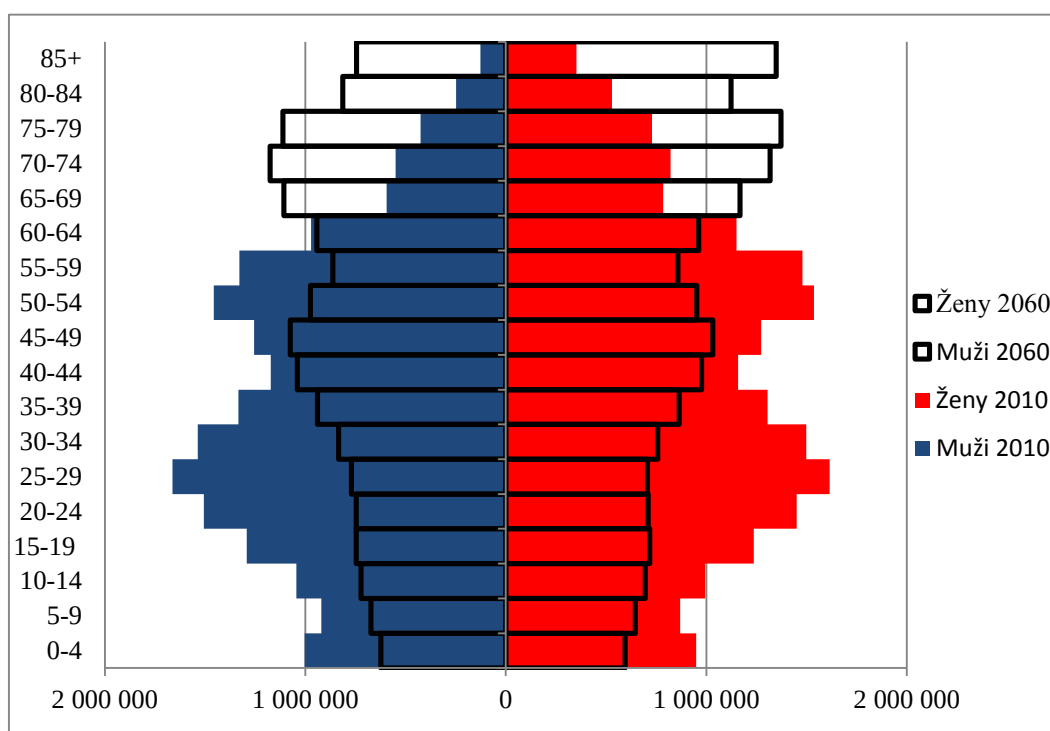
Obrázok 9: Porovnanie vekového zloženia obyvateľstva Maďarska v roku 2010 a 2060

Zdroj: Eurostat, vlastná konštrukcia



Obrázok 10: Porovnanie vekového zloženia obyvateľstva Slovenska v roku 2010 a 2060

Zdroj: Eurostat, vlastná konštrukcia



Obrázok 11: Porovnanie vekového zloženia obyvateľstva Poľska v roku 2010 a 2060

Zdroj: Eurostat, vlastná konštrukcia

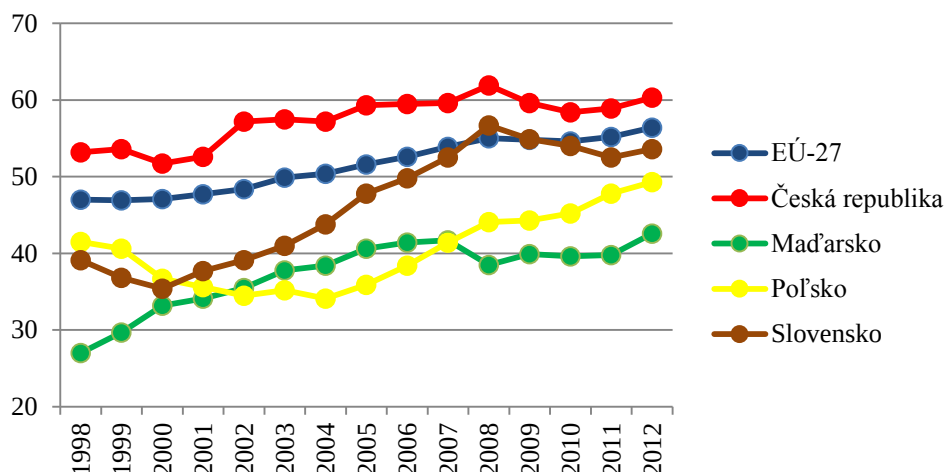
Demografickú dividendu pozorujeme pri európskych krajinách, vrátane krajín V-4 (viď obrázok 8, 9, 10, 11). Štádium vysokého podielu osôb v produktívnom veku znamená vyšší potenciál pri dosahovaní hospodárskeho rastu a zároveň nižšie hodnoty indexu závislosti. Problém nastane vtedy, keď sa obyvateľstvo v produktívnom veku dostane do poproduktívneho veku a zároveň detská zložka populácie bude

slabá. Tento vývoj potom spôsobuje populačné starnutie, prechod produktívnej časti populácie do starších vekových skupín a zvýšenie záťaže pri poskytovaní služieb závislým zložkám populácie (Slaný, 2007).

Z obrázku 8 a 9 vidíme, že Česká republika a Maďarsko disponujú vysokým podielom pracovnej sily a nízkym podielom detskej zložky obyvateľstva. Slovensko a Poľsko majú vyšší podiel detskej zložky populácie, ktorí by v budúcnosti predstavovali pracujúcu zložku obyvateľstva a vykompenzovali tak staršie osoby odchádzajúce do dôchodku (viď obrázok 10 a 11; Slaný, 2007).

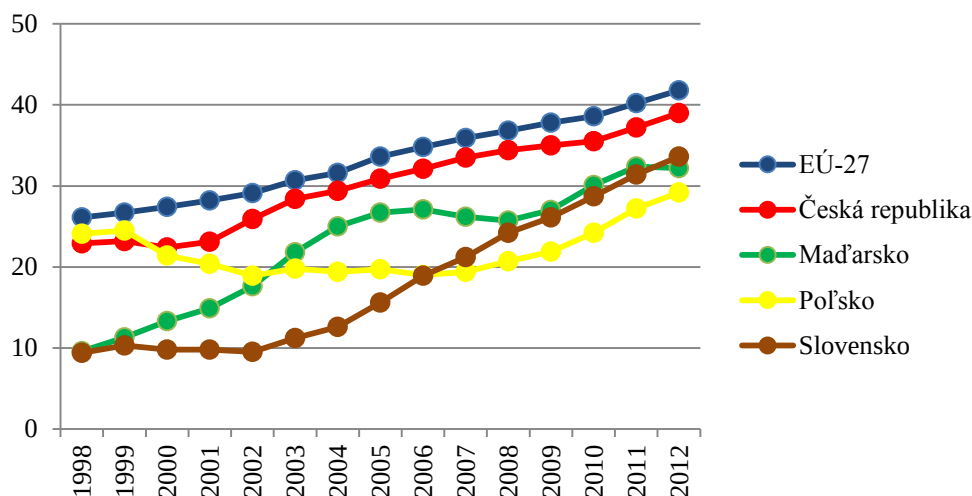
V otázke populačného starnutia je dôležitá a nevyhnutná zamestnanosť starších osôb a kontinuálne zvyšovanie dôchodkového veku. Zvýšenie veku odchodu do dôchodku pozitívne prispieva k väčšej zamestnanosti starších osôb (55–64) a predlžovanie aktívneho pracovného života kladne prispieva k udržaniu verejných financií. Z obrázkov 10 a 11 vyplýva, že vo všetkých sledovaných krajinách došlo k nárastu hodnôt mier zamestnanosti v období 1998–2012.

Česká republika patrí medzi krajiny, ktorých miera zamestnanosti mužov vo veku 55–64 sa nachádza nad priemerom členských krajín Európskej únie (viď obrázok 12). V sledovanom období sa zvýšila miera zamestnanosti mužov vo veku 55–64 rokov o 7,1%. Na druhej strane, miera zamestnanosti žien vo veku 55–64 rokov je v porovnaní s priemerom EÚ-27 nižšia (viď obrázok 13).



Obrázok 12: Miera zamestnanosti mužov vo veku 55–64 rokov v krajinách Vyšehradskej štvorky a EÚ-27 v období 1998–2012 (%)

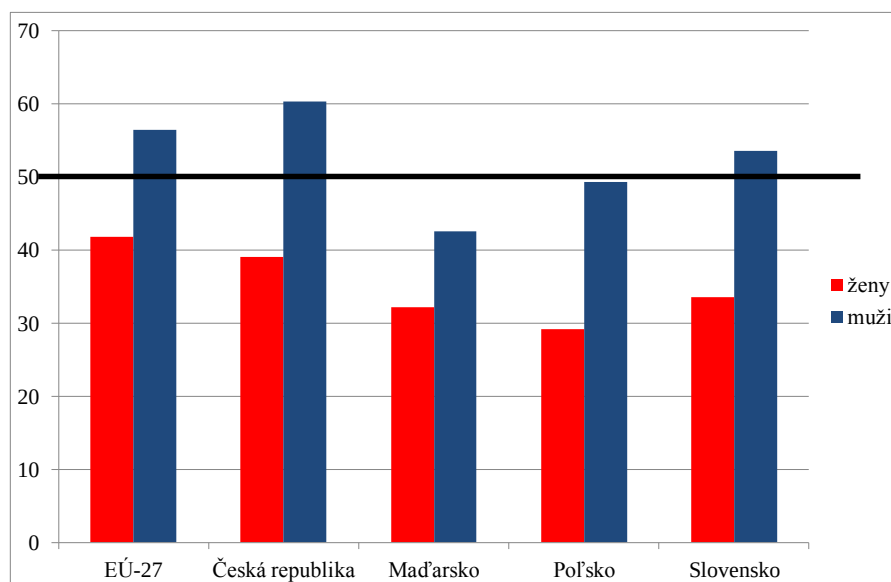
Zdroj: Eurostat, vlastná konštrukcia



Obrázok 13: Miera zamestnanosti žien vo veku 55–64 rokov v krajinách Vyšehradskej štvorky a EÚ-27 v období 1998–2012 (%)

Zdroj: Eurostat, vlastná konštrukcia

V prípade zamestnanosti staršej generácie je jednak dôležité ukončené vzdelanie a na druhej strane druh vykonávanej práce. V súčasnosti sa vyžadujú jazykové zručnosti a počítačové znalosti, ktorými potenciálna vekovo staršia pracovná sila mnohokrát nedisponuje. Cieľom je, aby sa zamestnanosť starších osôb zvýšila, ktorú významne ovplyvňuje ich zdravotný stav. Zastúpenie staršej pracovnej sily na trhu práce v sledovaných krajinách v porovnaní s priemerom krajín Európskej únie prezentuje obrázok 14. Vidíme, že miera zamestnanosti mužov vo veku 55–64 rokov v roku 2012 prekročila hranicu 50 % okrem EÚ-27 v Českej republike a na Slovensku, v ostatných krajinách, najmä v prípade žien, bola miera zamestnanosti výrazne nižšia.



Obrázok 14: Miera zamestnanosti mužov a žien vo veku 55–64 rokov v krajinách Vyšehradskej štvorky a EÚ-27 v roku 2012 (%)

Zdroj: Eurostat, vlastná konštrukcia

Záver

Populačné starnutie dopadá na všetky oblasti ekonomiky, preto by malo byť prioritným záujmom vládnych orgánov. Politická reakcia je oneskorená v porovnaní s demografickým vývojom. Na zmiernenie nepriaznivých účinkov demografického starnutia budú potrebné dlhodobé a premyslené politické postupy, pretože už nie je čas na ďalšie odďaľovanie tejto problematiky. Jedným riešením je propagácia myšlienky zamestnávania starších osôb firmami, ktoré by pomohli udržať produktivitu starších obyvateľov na požadovanej úrovni. Ďalšiu veľmi dôležitú úlohu bude zohrávať sociálna politika a prispôbenie sociálnych služieb starnúcej spoločnosti. Reforma penzijného systému v krajinách čeliacich populačnému starnutiu je nevyhnutná z hľadiska unesenia záťaže financovania systému.

Literatúra

Český statistický úrad. *Naděje dožití a průměrný věk: Metodika*. [online]. 2012 [cit. 2014-01-25]. Dostupné z: http://www.czso.cz/csu/redakce.nsf/i/nadeje_dozeni_a_prumerny vek

Důchodový věk v zemích EU. In: *Euroskop.cz* [online]. 2013 [cit. 2014-01-26]. Dostupné z: <https://www.euroskop.cz/578/sekce/duchodovy-vek-v-zemich-eu/>

Eurostat [online]. 2014 [cit. 2014-01-25]. Dostupné z: http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database

FN24: Hírek és szórakozás. *Kritikus idők – öregedő társadalom* [online]. 2005. [cit. 2014-01-25]. Dostupné z: http://fn.hir24.hu/csucsfigyaszto/2005/03/16/kritikus_id_337_k_oreged

LANGHAMROVÁ, Jitka. *Demografie: Učební text pro předmět U017* [online]. Praha, 2007 [cit. 2014-01-29]. ISBN 978-80-7399-218-7. Dostupné z: <http://u3v.vse.cz/wp-content/uploads/2009/03/U017.pdf>

LANGHAMROVÁ, Jitka. *Základy demografie: Materiály ke cvičením*. Praha: Oeconomica, 2013. ISBN 9788024519562.

SLANÝ, Antonín. *Faktory konkurenceschopnosti: Komparace zemí V – 4* [online]. Brno, 2007 [cit. 2014-01-29]. ISBN 978-80-210-4455-5. Dostupné z: https://is.muni.cz/do/econ/soubory/oddeleni/centrum/factory_konkurenceschopnosti.pdf

SZÉMAN, Zsuzsa. *Challenges of Ageing Societies in the Visegrad Countries: Hungary, Czech Republic, Poland, Slovakia* [online]. 2013 [cit. 2014-01-25]. ISBN 978-963-89445-4-2. Dostupné z: http://www.maltai.hu/data/Challenges_of_Aging_Societies.pdf

JEL Classification: J14

Summary

Demographic and Economic View of Population Ageing

The process of population aging is a highly discussed topic and it is the object of many debates. The process of demographic aging is associated with changes in the age structure of the population, while the most intensive increase is visible in the proportion of older people. Life expectancy is extending, while fertility rates are much lower. In today's advanced society, thanks to advances in medicine and health care, not only the increase of persons aged 60–70 will be visible, but the proportion of 80–90 years old or even 100 years and older will increase gradually. The changing age structure of the population of the Czech Republic will be a big challenge for the economy of the 21st century. In the future it will be necessary to pay a considerable amount of the government budget for social and health services. The character of the pension system needs to be radically changed. The work presents the development of selected demographic indicators related to population ageing in the Visegrad countries and economic problems associated with the aging society.

Keywords: Population Ageing, Age Structure of the Population, Old Age Dependency Ratio, Life Expectancy.

Jaký vliv budou mít modely používané pro extrapolaci křivky úmrtnosti na hodnotu střední a normální délky života?

Petra Dotlačilová

petra.dotlacilova@vse.cz

Doktorandka oboru statistika

Školitel: doc., Ing. Jitka Langhamrová, CSc. (langhamj@vse.cz)

Abstrakt: Článek se zabývá analýzou vývoje normální a střední délky života v České republice od roku 1920 do roku 2012. Cílem je porovnat vývoj obou zmíněných charakteristik u mužů a u žen. Z minulých výzkumů je zřejmé, že u všech vyspělých evropských populací dochází k postupnému zlepšování úmrtnosti. To znamená, že se postupně prodlužuje délka lidského života. Také se bude zvyšovat počet osob, které se dožívají nejvyšších věků.

Zlepšující se úmrtnost a prodlužující se délka lidského života způsobuje, že dříve nejpoužívanější analytická funkce (tj. Gompertzova-Makehamova funkce) přestává být nejvhodnější. Z tohoto důvodu bude v této studii představeno hned několik modelů, které se používají pro extrapolaci křivky úmrtnosti (např. Thatcherův model, Kannistův model, Coalův-Kiskerův model). Na základě těchto modelů bude nejprve vypočtena střední a potom normální délka života. Cílem je zjistit, jak se bude lišit normální délka života při použití vybraných modelů. Výsledky budou také porovnány s hodnotami střední délky života při narození.

Získané výstupy pro normální délku života a střední délku života při narození budou publikovány podle jednotlivých modelů a zároveň budou publikovány zvlášť pro muže a ženy.

Klíčová slova: úmrtnost, střední délka života, normální délka života.

Úvod

Úmrtnost tvoří spolu s plodností základní složku reprodukce obyvatelstva. Demografové analyzují závislost úmrtnosti na věku i na pohlaví. Ze současných vědeckých výzkumů plyne, že dochází k neustálému snižování úmrtnosti. Což znamená, že se prodlužuje délka lidského života. Tento vývoj je způsobem především procesem modernizace a neustálými pokroky v medicíně, ale i propagací zdravého životního stylu, zvyšováním bezpečnosti práce i dopravních prostředků atd.

Pro popis úmrtnosti je možné použít několik ukazatelů. Nejčastěji se používá syntetický ukazatel známý jako střední délka života. Jako další můžeme zmínit normální nebo pravděpodobnou délku života.

Při zkoumání úmrtnosti v jednotlivých věcích je však třeba uvědomit si, že úmrtnost u nejstarších osob vykazuje jiný vývoj než úmrtnost u mladších osob. Má více biologický charakter a méně charakter sociální, výrazněji závisí na věku. Ve velmi vysokém věku, kde jsou počty žijících malé, narážíme na problém výrazného náhodného kolísání empirických hodnot specifických měr úmrtnosti. Proto v tomto věku úmrtnost modelujeme nějakou analytickou funkcí. V minulosti se nejvíce využívala Gompertzova-Makehamova funkce. Dnes se vzhledem k měnícímu se vývoji úmrtnosti nabízí použití ještě několika dalších dostupných modelů.

V této studii budou vybrané modely použity pro extrapolaci křivky úmrtnosti. Pomocí získaných výstupů bude vypočtena střední a normální délka života.

Metodika

Jak již bylo řečeno, k popisu úmrtnosti se nejčastěji používá ukazatel známý jako střední délka života. Není to však jediný ukazatel. Jako další můžeme zmínit normální nebo pravděpodobnou délku života.

Pro získání střední délky života je potřeba nejprve vypočítat úmrtnostní tabulky. Jejich výpočet provedeme na základě dat o počtech zemřelých a žijících (rozlišených podle věku a pohlaví) a počtech živě narozených. Nejprve vypočteme specifické míry úmrtnosti, které získáme jako počet zemřelých v dokončeném věku x vztahených ke střednímu stavu počtu žijících v daném věku. Pro usnadnění analýzy

je vhodné nahradit diskrétní model úmrtnosti (založený na specifických mírách úmrtnosti v jednoletých nebo pětiletých intervalech) modelem spojitým a definovat bodovou charakteristiku úmrtnosti – intenzitu úmrtnosti.

V mladších věcích (kde lze předpokládat, že v jednoletých intervalech je vývoj intenzity úmrtnosti prakticky lineární) lze hodnoty intenzity úmrtnosti odhadnout na základě vztahu (Fiala, 2005):

$$m_x = \mu(x + 0,5), x = 1, 2, \dots, 59, \quad (1)$$

kde m_x je specifická míra úmrtnosti a $\mu(x + 0,5)$ je intenzita úmrtnosti ve věku $x + 0,5$.

Pro vyšší věky je intenzita úmrtnosti popisována některým z dostupných modelů. K dříve nepoužívanějším modelům patří již dříve zmíněná Gompertzova-Makehamova funkce. Jako další modely budou použity: Kannistův model, Thatcherův model, Heligmanův-Pollardův model, Coalův-Kiskerův model (Burcin et al., 2010).

Gompertzova-Makehamova funkce se stala na dlouhou dobu nepoužívanější funkcí používanou pro extrapolaci křivky úmrtnosti. Tato funkce předpokládá, že dochází k růstu úmrtnosti a to exponenciálně s rostoucím věkem. Při dalších výzkumech bylo zjištěno, že tento předpoklad není splněn pro nejvyšší věky. To způsobuje, že s prodlužující se délkou lidského života se tato funkce stává méně vhodnou (Gompertz, 1825).

Jako další můžeme uvést model Coala-Kiskera. Ten se soustředí na změny měr úmrtnosti mezi dvěma po sobě jdoucími věky. Vývoj úmrtnosti lze popsat exponenciálně kvadratickou funkcí.

Následující dva modely patří do skupiny logistických modelů. V této studii bude uveden Thatcherův a Kannistův model. U Kannistova modelu autor zjistil, že použití tohoto modelu je vhodné v nejvyšších věcích. Pro nejnovější data o μ_x v nejvyšších věcích platí, že odpovídají logistickému modelu, ve kterém je $\logit(\mu_x)$ vyjádřen jako lineární funkce věku (x). Druhým logistickým modelem je Thatcherův model, který vychází z předpokladu logistického průběhu křivky úmrtnosti (Thatcher et al., 1998).

Je zřejmé, že při použití logistických modelů budeme dosahovat vyšších hodnot střední délky života. Oba modely tedy patří k optimističtějším.

Jako další bude zmíněn Heligmanův-Pollardův model. Ten se z hlediska dosažených hodnot střední délky života neřadí ani mezi optimistické, ani mezi pesimistické modely a vyskytuje se někde mezi hodnotami středních délek života získaných podle již výše uvedených modelů. Model v původní formě obsahoval osm parametrů. Pokud se zaměříme jen na modelování úmrtnosti ve vyšších věcích, tak bude mít model uvedenou podobu (viz příloha 1).

Podrobnější popis jednotlivých modelů je uveden v příloze 1.

Jako druhý ukazatel bude použita normální délka života. Definujeme ji jako věk, ve kterém se nejvíce umírá. Její výpočet provedeme podle vztahu (Fiala, 2005):

$$\hat{e}_x = \max(d_x) + 0,5, \quad (2)$$

kde d_x je počet zemřelých z tabulkové populace v dokončeném věku x .

Na základě uvedeného vztahu bude proveden odhad hodnoty normální délky pro všechny zmíněné modely.

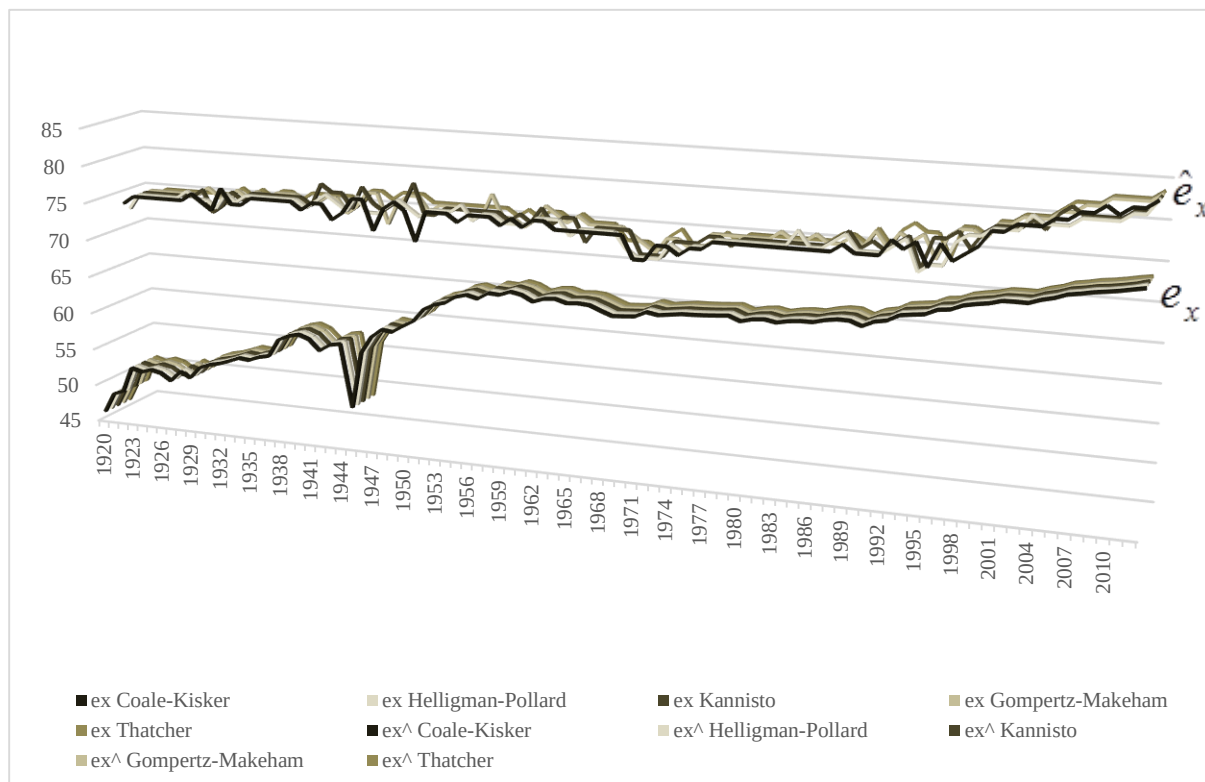
Při výpočtech je důležité uvědomit si, že střední délka života je charakteristika typu průměr a normální délka života je charakteristika typu modus. Nevýhodou střední délky života je, že může být ovlivněna extrémními hodnotami (především vysokou kojeneckou úmrtností). Oproti tomu normální délka života je definována jako věk, ve kterém zemře nejvíce osob. Je to tedy charakteristika, která není ovlivněna extrémními hodnotami. Jde o ukazatel, který se velmi často používá jako doplňkový ukazatel ke střední délce života.

Dosažené výsledky

Cílem tohoto příspěvku je provedení odhadu střední a normální délky života při použití vybraných modelů používaných pro extrapolaci křivky úmrtnosti. Dále zjistit jaké jsou rozdíly mezi hodnotami normální délky života při aplikaci vybraných modelů a získané výsledky porovnat s hodnotami střední délky života u narozených osob.

Jako vstupní data byla použita data o úmrtnosti české populace od roku 1920 do roku 2012. Získané výsledky budou publikovány jak pro muže, tak pro ženy.

Na následujícím grafu je vyobrazen vývoj střední délky života při narození a normální délky života získané při aplikaci modelů Coale-Kiskera, Heligmana-Pollarda, Thatcher, Kannista a Gompertzovy-Makehamovy funkce.



Graf 1 – Vývoj střední a normální délky života – Česká republika - muži¹² Zdroj: data ČSÚ, výpočty vlastní

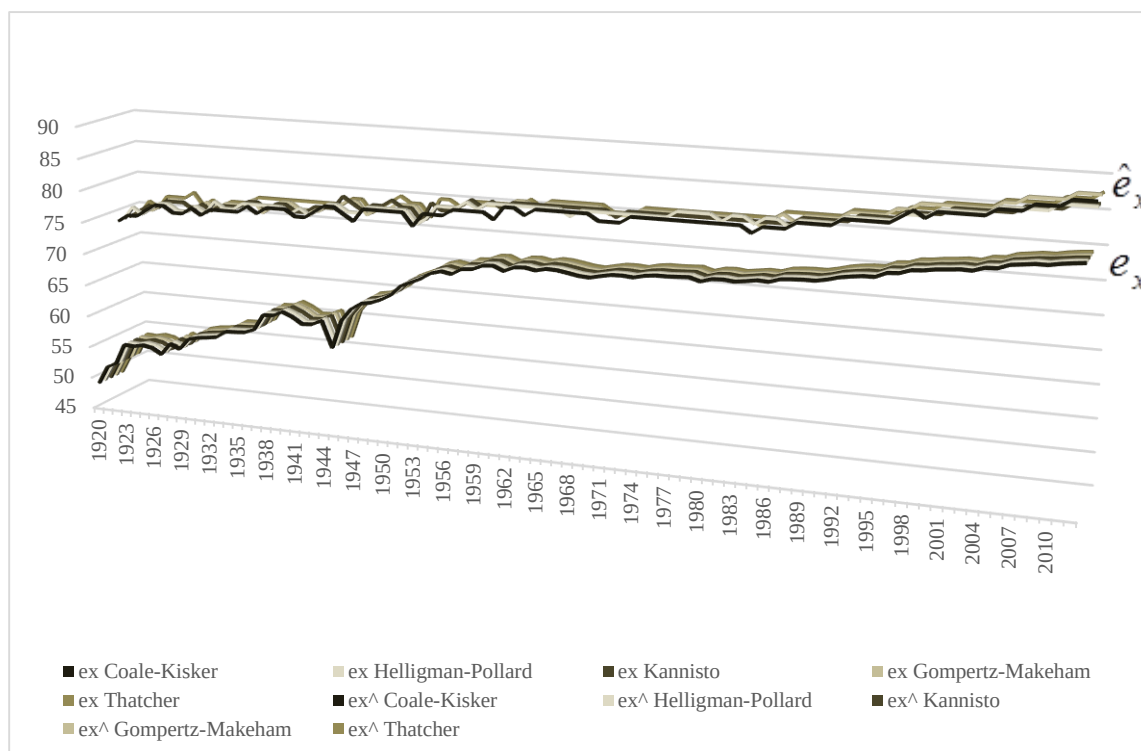
Z grafu 1 je patrné, že v průběhu celého sledovaného období došlo k růstu obou sledovaných ukazatelů. K významnějšímu nárůstu došlo u střední délky života při narození. To je způsobeno především snížením kojenecké úmrtnosti. Výrazný pokles pozorujeme v roce 1945. Pokles je způsoben především zvýšením počtu chlapců zemřelých v dokončeném věku 0 a zároveň snížením počtu živě narozených chlapců, ke kterému došlo v roce 1945. V následujícím roce opět pokračuje nárůst. Během 60–80. let však střední délka života mužů v ČR prakticky stagnovala, k obnovení nárůstu došlo teprve od roku 1990. V průběhu celého sledovaného období vzrostla hodnota sledovaného ukazatele o 26,2 roku (tj. z hodnoty 48,8 na hodnotu 75). Pokud budeme porovnávat získané výsledky získané při aplikaci vybraných modelů, tak zjišťujeme, že mezi získanými hodnotami jsou pouze minimální rozdíly (rozdíly se objevují na místě setin). Malé rozdíly v hodnotách střední délky života při narození (podle jednotlivých modelů) jsou způsobeny jejich malým dopadem na hodnotu tohoto ukazatele u právě narozených osob (důležitost použití jednotlivých modelů je patrná především v nejvyšších věcích).

¹ Symbolem e_x je v grafu označena střední délka života při narození a symbolem $e_x^{\wedge}$ je označena normální délka života.

² Pro výpočty byl použit software DeRaS (Burcin et al., 2012).

Pokud se zaměříme na vývoj normální délky života, tak zjišťujeme, že v průběhu sledovaného období nedošlo k tak velkému nárůstu. To je způsobeno především tím, že normální délka života je charakteristikou typu modus a tudíž její hodnota není ovlivněna extrémními hodnotami. Na hodnotu tohoto ukazatele tedy nemají vliv změny v kojenecké úmrtnosti. V průběhu sledovaného období došlo k nárůstu normální délky života o 9 let (z hodnoty 74,5 na hodnotu 83,5). Pokud budeme porovnávat vývoj v celém sledovaném období podle jednotlivých modelů, dojdeme k závěru, že hodnoty normální délky života jsou v průběhu celého sledovaného období rozkolísané. To je způsobeno především tím, že odhad normální délky života byl proveden pomocí (2), což je pouze hrubý odhad hodnoty normální délky života. Při porovnání výsledků podle vybraných modelů zjistíme, že nejvyšší hodnoty poskytuje Thatcherův model. Naopak nejnižší hodnoty dostaneme při použití Coalova-Kiskerova modelu.

Na druhém grafu je znázorněn vývoj střední a normální délky života od roku 1920 do roku 2012 pro ženy v České republice získané při aplikaci vybraných modelů.



Graf 2 – Vývoj střední a normální délky života – Česká republika – ženy Zdroj: data ČSÚ, výpočty vlastní

Z grafu 2 je zřejmé, že v průběhu sledovaného období došlo k nárůstu obou sledovaných charakteristik (stejně jako u mužů). Střední délka života při narození vzrostla přibližně o 29,1 let (z 51,8 na 80,9 let). Je tedy zřejmé, že u žen došlo k většímu nárůstu střední délky života. Jedním z důvodů může být, že u žen nedošlo k tak významnému poklesu v roce 1945. V posledním válečném roce poklesla hodnota střední délky života při narození o 4,1 roku. Pokud se blíže zaměříme na výsledky získané při aplikaci jednotlivých modelů, tak docházíme ke stejnému závěru jako u mužů. Při bližším zkoumání zjišťujeme, že rozdíly v hodnotách střední délky života (podle vybraných modelů) se projevují až na místě setin. Malé rozdíly v hodnotách střední délky života při narození (podle jednotlivých modelů) jsou způsobeny malým dopadem na hodnotu tohoto ukazatele u právě narozených osob (důležitost použití jednotlivých modelů je patrná především v nejvyšších věcích).

Při zkoumání vývoje normální délky života je zřejmé, že i zde došlo k nárůstu. Z grafu 2 je také patrné, že v průběhu celého sledovaného období nedocházelo k významnějším výkyvům. Za celé sledované období vzrostla normální délka života u českých žen o 14 let (z hodnoty 73,5 na 87,5). Pokud budeme porovnávat získané výsledky při aplikaci jednotlivých modelů, zjistíme, že nejvyšší hodnoty získáme při aplikaci Thatcherova modelu. Naopak nejnižší hodnoty získáváme u Coalova-Kiskerova modelu (zhruba do roku 2002). Po roce 2002 jsou nižší hodnoty získány například u Helligmanova-Pollardova modelu.

Závěr

Ze získaných výsledků je zřejmé, že v průběhu celého sledovaného období došlo k nárůstu obou charakteristik délky života. Jak u mužů, tak u žen došlo k většímu nárůstu u střední délky života (v porovnání s normální délkou života). Příčinou tohoto výraznějšího nárůstu je především snížení kojenecké úmrtnosti, ke kterému došlo v průběhu sledovaného období. Pokud budeme porovnávat hodnoty získané při aplikaci jednotlivých modelů je zřejmé, že se získané hodnoty od sebe liší jen nepatrně (rozdíly se objevují v řádech setin).

Pokud se zaměříme na zkoumání vývoje normální délky života, docházíme k závěru, že i zde docházelo v průběhu celého období k nárůstu. Také je patrné, že nárůst nebyl tak významný, a také docházelo k významnějšímu kolísání hodnoty tohoto ukazatele. Pokud budeme porovnávat výsledky podle jednotlivých modelů, zjišťujeme, že nejvyšší rozdíl mezi získanými hodnotami je jeden rok. Dále je zřejmé, že k neoptimističtějším modelům patří Thatcherův model a naopak mezi ty pesimistické patří model Coala-Kiskera.

Jak již bylo zmíněno v úvodu, cílem příspěvku bylo posoudit, jak ovlivní použití jednotlivých modelů hodnotu normální délky života. Pro porovnání je vše doplněno hodnotami střední délky života při narození (získané při aplikaci jednotlivých modelů). Na základě získaných výsledků můžeme usoudit, že (stejně jako u hodnot střední délky života) nejvyšší při aplikaci Thatcherova modelu. Pokud budeme zjišťovat, který z modelů nejlépe popisuje závislost intenzity úmrtnosti na věku, tak zjistíme, že nelze vyslovit jednoznačný závěr. U žen se pro popis uvedené závislosti nejvíce hodí Heligmanův-Pollardův nebo Coalův-Kiskerův model (platí přibližně do 70. let). Později se jako vhodnější ukazují modely logistické Thatcherův nebo Kannistův. U mužů je situace trochu odlišná. Do roku 1950 se jako nejvhodnější ukazuje model Heligmana-Pollarda. Od roku 1950 lépe vychází Coalův-Kiskerův model (pro některé vybrané roky ještě lépe vychází Gompertzova-Makehamova funkce). Od roku 2002 se pro popis závislosti intenzity úmrtnosti na věku nejlépe hodí logistické modely (Thatcherův nebo Kannistův).

Na závěr je důležité uvést, že vývoj obou zkoumaných ukazatelů lze také zdůvodnit tím, o jaké typy ukazatelů se jedná. Střední délka života je charakteristikou typu průměr a její hodnota je tedy ovlivněna extrémními hodnotami. Proto je vhodné kromě tohoto ukazatele použít ještě jiný, který není ovlivněn odlehými pozorováními (např. normální délka života).

Příloha 1

Autoři	Model	Parametry modelu	Věkové rozpětí pro vyrovnání vybraným modelem
Gompertz-Makeham	$\mu_{x+0,5} = a + bc^{x+0,5}$	a, b, c	65-90
Coale-Kisker	$m_x = e^{ax^2 + bx + c}$	a, b, c	65-90
Thatcher	$\mu_{x+0,5} = c + \frac{ae^{b(x+0,5)}}{1 + ae^{b(x+0,5)}}$	a, b, c	65-90
Kannistö	$\mu_{x+0,5} = \frac{ae^{b(x+0,5)}}{1 + ae^{b(x+0,5)}}$	a, b	65-90
Heligman-Pollard	$q_x = \frac{ae^{bx}}{1 + ae^{bx}}$	a, b	65-90

Zdroj: Thatcher et al., 1998

Příloha 2

Model	Hodnoty střední délky života při narození	Hodnoty normální délky života
Gompertz-Makeham	nižší hodnoty	hodnoty se vyskytují mezi ostatními modely
Coale-Kisker	nižší hodnoty	nejnižší hodnoty (do roku 1992)
Thatcher	vyšší hodnoty	nejvyšší hodnoty
Kannistö	vyšší hodnoty	vyšší hodnoty
Heligman-Pollard	blíží se Thatcherovu modelu	nejnižší hodnoty (od roku 1992)

Literatura

BURCIN, B., TESÁRKOVÁ, K. a ŠÍDLO, L.: Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci. *Demografie*. 2010, č. 52, s. 77 – 89.

BURCIN, B., HULÍKOVÁ TESÁRKOVÁ, K. a KOMÁNEK, D. (2012): DeRaS: software tool for modelling mortality intensities and life table construction. Charles University in Prague, Prague. <http://deras.natur.cuni.cz>.

CIPRA, T.: Matematické metody demografie a pojištění. 1990. Praha, STNL.

ČSÚ, 2014. Dostupné z WWW: http://www.czso.cz/csu/redakce.nsf/i/obyvatelstvo_lide.

DOTLAČILOVÁ, P.: Rozdíly v úmrtnosti české a slovenské populace. In: *Vědecká konference doktorského studia FIS 2013 [CD-ROM]*. Praha, 14.02.2013. Praha: Oeconomica, 2013, s. 174–181. ISBN 978-80-245-1924-1.

FIALA, T.: *Výpočty aktuárské demografie v tabulkovém procesoru*. Praha: Vysoká škola ekonomická v Praze, 2005. ISBN 80-2450821-4.

GOMPERTZ, B.: On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. *Philosophical Transactions of the Royal Society of London* 115 (1825): 513–585.

KOSCHIN, Felix. *Aktuárská demografie*. 1. vyd. Praha: VŠE, 2002. 100 s. ISBN 80-245-0403-0.

MAKEHAM, W., M.: On the Law of Mortality and the Construction of Annuity Tables. *The Assurance Magazine, and Journal of the Institute of Actuaries* 8 (1860): 301–310.

THATCHER, R., A., KANISTÖ, V. a VAUPEL, J., W.: *The Force of Mortality at Ages 80 to 120*. 1998. ISBN 87-7838-381-1.

JEL Classification: C02, C10

Summary

How will be the influence of models used for extrapolation mortality curves on the value of life expectancy and modal age at death?

The study analyzes the differences in values of life expectancy at birth and modal age at death when we applied selected models that are used for extrapolation of mortality curves. As the input data will be used the data about mortality of the Czech population from 1920 to 2012. Each model will be applied on this data and then will be obtained values of life expectancy and modal age at death. This values will be compared for each model and for the entire reporting period.

Key words: mortality, life expectancy, modal age at death.

Nadúmrtnost mužů v ČR

Jana Langhamrová

Xlanj18@vse.cz

Doktorand oboru statistika

Školitel: doc. Ing. Markéta Arltová, Ph.D., (arltova@vse.cz)

Abstrakt: Cílem příspěvku bylo popsat vývoj vybraných ukazatelů úmrtnosti v České republice v letech 1950–2012. Na základě údajů Českého statistického úřadu o počtech zemřelých mužů a žen a věkové struktuře mužů a žen byly zkonstruovány úmrtnostní tabulky. A na jejich základě vypočteny délky života v přesných věcích. Příspěvek poukazuje na problematiku tzv. mužské nadúmrtnosti. Pro muže a ženy byly vypočítány kojenecké úmrtnosti. Na základě rozboru a srovnání vývoje těchto úmrtností se zde poukazuje na fakt, že i nejnižším věku mají muži vyšší intenzitu úmrtnosti než ženy. Mužskou nadúmrtnost je možné zkoumat porovnáním středních délek života žen a mužů v celém sledovaném období. Do budoucna lze předpokládat, že se střední délka života bude jak pro muže, tak pro ženy dále prodlužovat. Toto prodloužení nebude však již z důvodu klesající kojenecké úmrtnosti (ta je již na velmi nízké a světově srovnatelné úrovni), ale z důvodu poklesu a zlepšení úmrtnosti ve středních a vyšších věcích. Rozdíly ve střední délce života žen a mužů se pak budou v důsledku těchto změn dále mírně snižovat.

Klíčová slova: mužská nadúmrtnost, střední délka života, naděje dožití, kojenecká úmrtnost, kvocient kojenecké úmrtnosti.

Úvod

Všeobecně známým faktem je, že ve vyspělých zemích je možné pozorovat tzv. mužskou nadúmrtnost. Mužská nadúmrtnost má za následek skutečnost, že i když se rodí více chlapečků než holčiček (pravděpodobnost narození chlapce je okolo 0,515), ve starším věku převládají ženy.

Proč existují rozdíly mezi mužskou a ženskou úmrtností? Tomuto tématu se věnují nejen lékaři. Uvažuje se, že pokud by neexistovaly rušivé vlivy, které zkracují život mužů, měl by být rozdíl mezi střední délkou života muže a ženy při narození pouze 1 rok v neprospěch mužů. Střední délka života nebo také naděje dožití je jedním z ukazatelů úmrtnosti, který bývá používán jako určitý indikátor vyspělosti dané společnosti nebo také zdravotního stavu populace. Zjednodušeně můžeme říci, že jde o odhad průměrného počtu let, které má ještě před sebou přesně x -letá osoba, pokud budou platit po zbytek jejího života stávající úmrtnostní poměry.

V tomto příspěvku bude zkoumán vliv mužské nadúmrtnosti v letech 1950–2012 na již zmíněnou střední délku života žen a mužů při narození a ve věku 65 let. Pozornost bude také věnována kvocientu kojenecké úmrtnosti podle pohlaví. Pomocí metody dekompozice budou vypočteny podíly věkových skupin na celkovém rozdílu v délce života žen a mužů.

Rozdíly v úmrtnosti mužů a žen – historie a současnost

Na počátku 20. let minulého století byl rozdíl mezi střední délkou života mužů a žen při narození okolo 2,5 let a postupně tento rozdíl narůstal. Na počátku 60. let byl rozdíl mezi muži a ženami 5,5 roku a na počátku 90. let dosáhl téměř 8 let. Teprve v současnosti se stabilizuje na hodnotě okolo 6 let.

Po průmyslové revoluci a v období tzv. modernizace společnosti významně klesá úmrtnost. Již Talacko (1941) upozornil na pokles úmrtnosti mužů a žen v 2. polovině 19. století. Poukazuje na skutečnost, že od 70. let 19. století do roku 1900 klesá úmrtnost. Největší pokles byl pozorován u obou pohlaví v raném dětském věku. Dále uvádí, že „*ve všech stářích nad 40 roků poklesla úmrtnost u žen více než u mužů... v kojeneckém věku je úmrtnost žen vždy menší než mužů, v raném dětském věku rozdíl není podstatný*“. Nejpodstatnější rozdíly z období 2. poloviny 19. století mezi ženami a muži lze pozorovat v době dospívání, v době největší plodnosti ženy (to již dnes neplatí, není již vyšší úmrtnost žen, způsobená mateřskou úmrtností) a v době od 40–45 roků života. Talacko si povšiml, že „*úmrtnost mužů nad 40 roků je u nás abnormálně vysoká*“ (myšleno ve srovnání s ostatními zeměmi). Již zde je tedy poukázáno na

vliv určitých aspektů, které zapříčinily vyšší úmrtnost mužů. Můžeme hovořit o určité inklinaci mužů k některým chorobám, otázkou také je, zda především v dřívější době nemá vliv na mužskou nadúmrt-nost těžší práce mužů a zaměstnání v horších či nebezpečnějších pracovních podmínkách.

Současným problémem většiny vyspělých zemí je vysoká úmrtnost na nemoci oběhové soustavy. V časopise *Medicína* se uvádí: „Výrazné zvýšení výskytu ischemické choroby srdeční u mužů, o kterém se mluví jako o epidemii, začalo ve Spojených státech počátkem 20. let minulého století, ve Velké Británii a dalších evropských průmyslových zemích krátce po druhé světové válce. Mortalita mužů při ní byla zřetelně vyšší. Podle převládajícího mínění je příčinou tohoto rozdílu mezi pohlavími protektivní působení estrogenu u žen. Ale klinické studie ze šedesátých a sedmdesátých let, které ověřovaly možnost prevence ICHS u mužů vysokými dávkami estrogenu, skončily neúspěchem.“ [Medicína, 2001]

Proto také byly na univerzitě v Bristolu analyzovány dostupné údaje z let 1921–1998 o úmrtnosti obyvatel Velké Británie s cílem popsat dlouhodobé trendy v úmrtnosti a najít příčinný faktor vyšší úmrtnosti mužů. Autoři výsledky srovnávali s ostatními zeměmi a došli k zajímavému závěru. Pokud se zaměřili na velmi častou příčinu smrti – ischemickou chorobu srdeční – tak zvýšená úmrtnost mužů není způsobena tím, že nejsou chráněni působením estrogenu (ten chrání ženy a proto ženy mají vyšší výskyt úmrtnosti na tuto chorobu až v době po přechodu). Protože zkoumali více zemí a více generací žijících ve stejném období, hledali faktor prostředí, který by byl společný. Zkoumali korelaci mezi zvýšenou úmrtností mužů a řadou dalších faktorů. Významný se jevil vztah mezi úmrtností a zvýšenou konzumací tuků. Lze se také domnívat, že významný vliv má také způsob života a geografická poloha. Co je však zajímavé, že tato studie nenašla souvislost mezi vyšší úmrtností a kuřáctvím ani vyšší konzumací alkoholu.

Jiný názor na alkohol a tabák je popsán v článku McCartney a kol. (2011). Zde se naopak setkáváme s názorem, že kouření a alkohol mají vliv na rozdíl v úmrtnosti mužů a žen. Nepřiměřená konzumace alkoholu má jistě vliv na úmrtnost. V této souvislosti se hovoří také o tzv. ruské mortalitní krizi, kdy lze pozorovat, jak se mění střední délka života především mužů v souvislosti s obdobími zpřísněných opatření umožňujících požívat alkohol. I zde se však popírá fakt, že ženy jsou biologicky lépe vybaveny. Uvádí se skutečnost, že se liší úmrtnost podle pohlaví v jednotlivých zemích. Zdůrazněn je fakt, že úmrtnost a rozdíly mezi pohlavími jsou dány také sociálním prostředím.

Již delší dobu se vědci snaží vysvětlit, proč dochází k nadúmrti mužů, která je v současné době typická nejenom pro Českou republiku, ale také většinu zemí světa.

Obecně lze říci, že muži jsou z biologického pohledu hůře vybaveni, mají nižší biologickou imunitu než ženy. Jak již bylo zmíněno dříve, roli hrají také sociokulturní vlivy. Muži a ženy mají ve společnosti různou úlohu a s ní spojenou zodpovědnost. Velký vliv má životní styl člověka, muži jsou častěji vystavováni rizikovým faktorům (horší pracovní prostředí, těžší fyzická práce, větší podíl kuřáků, vyšší konzumace alkoholu, horší stravovací návyky). Ženy přistupují ke svému zdraví zodpovědněji. Zajímavou otázkou, kterou se autoři těchto studií zabývají, může být, zda se neprojeví do budoucna v intenzitě úmrtnosti žen fakt, že některé z nich přejímají tzv. mužské vzory chování.

Výpočet vybraných ukazatelů a charakteristik

Výchozím údajem, který je v tomto příspěvku využit pro výpočet ukazatelů úmrtnosti, je absolutní počet zemřelých, za určité časové období (zpravidla rok, měsíc) na sledovaném území. Počet zemřelých ovšem není vhodnou charakteristikou, protože nezohledňuje velikost populace, ve které je počet zemřelých sledován. Velmi často se proto vypočítává tzv. obecná míra úmrtnosti nebo také hrubá míra úmrtnosti, která tento problém částečně řeší. Jde o poměr počtu zemřelých osob a středního stavu obyvatel (stav obyvatel k 1. 7. daného roku). Obecná míra úmrtnosti závisí na věkové struktuře obyvatelstva a intenzitách úmrtnosti v jednotlivých věkových skupinách. Její výsledek je závislý nejenom na samotné intenzitě úmrtnosti podle věku, ale také na příznivé či nepříznivé věkové struktuře obyvatelstva.

Rozdíly v úmrtnosti podle pohlaví lze tedy jednoduše vyjádřit porovnáním specifických měr a výpočtem indexu, který porovnává intenzitu úmrtnosti žen a mužů. Specifická míra úmrtnosti se značí $m_{t,x}$ a vypočítá se jako počet zemřelých osob v dokončeném věku x let na 1 000 obyvatel ve věku x let v příslušném roce t :

$$m_{t,x} = \frac{M_{t,x}}{S_{t,x}},$$

kde

$M_{t,x}$ – počet zemřelých ve věku dokončených x let v roce t ,

$S_{t,x}$ – střední stav obyvatelstva ve věku dokončených x let v roce t .

Víme, že úmrtnost a její intenzita se s věkem významně mění. Demografové se proto často zaměřují na úmrtnost na počátku života. Sleduje se především kojenecká úmrtnost (kvocient kojenecké úmrtnosti). Největší riziko úmrtí je pro člověka těsně po narození a následně v období vysokého věku. Proto se věnuje značná pozornost právě kojenecké úmrtnosti. Vypočítá se nejčastěji jako počet zemřelých dětí do jednoho roku života na 1 000 živě narozených dětí ve stejném kalendářním roce:

$$m_{t,0} = \frac{M_{t,0}}{N_t},$$

kde

$M_{t,0}$ – počet zemřelých ve věku 0 let v roce t ,

N_t – počet narozených v roce t .

Jedná se o jeden z nejdůležitějších ukazatelů, který vystihuje zdravotní stav populace, zdravotnictví a životní úroveň dané země. Protože je úmrtnost v prvním roce života dítěte značně nerovnoměrná, počítají se další ukazatele, například kvocient úmrtnosti prvního dne, poporodní úmrtnost (úmrtí do 2 dnů), časná novorozenecká úmrtnost (úmrtí do 6 dnů), novorozenecká úmrtnost (úmrtí do 4 týdnů) nebo kvocientem ponovorozenecké úmrtnosti (úmrtí v období od 4 týdnů do 1 roku). Pro účely tohoto příspěvku byly vybrány ukazatele kojenecké úmrtnosti vypočteny zvlášť pro chlapce a dívky.

Chceme-li porovnávat, jak se liší úmrtnost mužů a žen, nejčastěji se setkáváme s porovnáním středních délek života v přesném věku x let. Obvykle se používá střední délka života při narození a střední délka života v přesném věku 65 let. Při konstrukci těchto ukazatelů se vychází z modelu úmrtnostních tabulek (podrobný postup výpočtu viz Roubíček, 1997; Fiala, 2005). Vzhledem k rozsahu příspěvku zde nebudeme uvádět všechny obvyklé způsoby výpočtu úmrtnostních tabulek a modelů vyrovnávání úmrtnosti ve starších věcích. V tomto příspěvku byl pro výpočet úmrtnostních tabulek a střední délky života zvolen postup podle Fiala (2005). Podrobnější postup výpočtu lze nalézt v Langhamrová (2013). Pro vyšší věky bylo provedeno vyrovnání pomocí Gompertzovy-Makehamovy funkce. Střední délku života osoby v přesném věku x lze popsat jako poměr počtu zbylých let života T_x a počtu dožívajících se přesného věku l_x :

$$e_x^0 = \frac{T_x}{l_x},$$

kde

T_x – počet zbylých let života,

l_x – počet osob dožívajících se přesného věku.

Rozdíl ve střední délce života žen a mužů lze hodnotit také pomocí tzv. metody dekompozice. Pokud dojde např. mezi lety k nárůstu střední délky života, lze tento nárůst rozebrat a zjistit, např., které věkové skupiny se na tomto nárůstu střední délky života podílely. Příspěvky jednotlivých věkových skupin ke změně střední délky života je možné počítat pomocí jednorozměrné dekompozice Pressat (1972):

$$\Delta_{x,x+n} = (e_x^A - e_x^B) * \frac{l_x^A + l_x^B}{200000} - (e_{x+n}^A - e_{x+n}^B) * \frac{l_{x+n}^A + l_{x+n}^B}{200000},$$

kde

e_x, e_{x+n} – střední délka života ve věku x , resp. $x + n$ v roce A, B ,

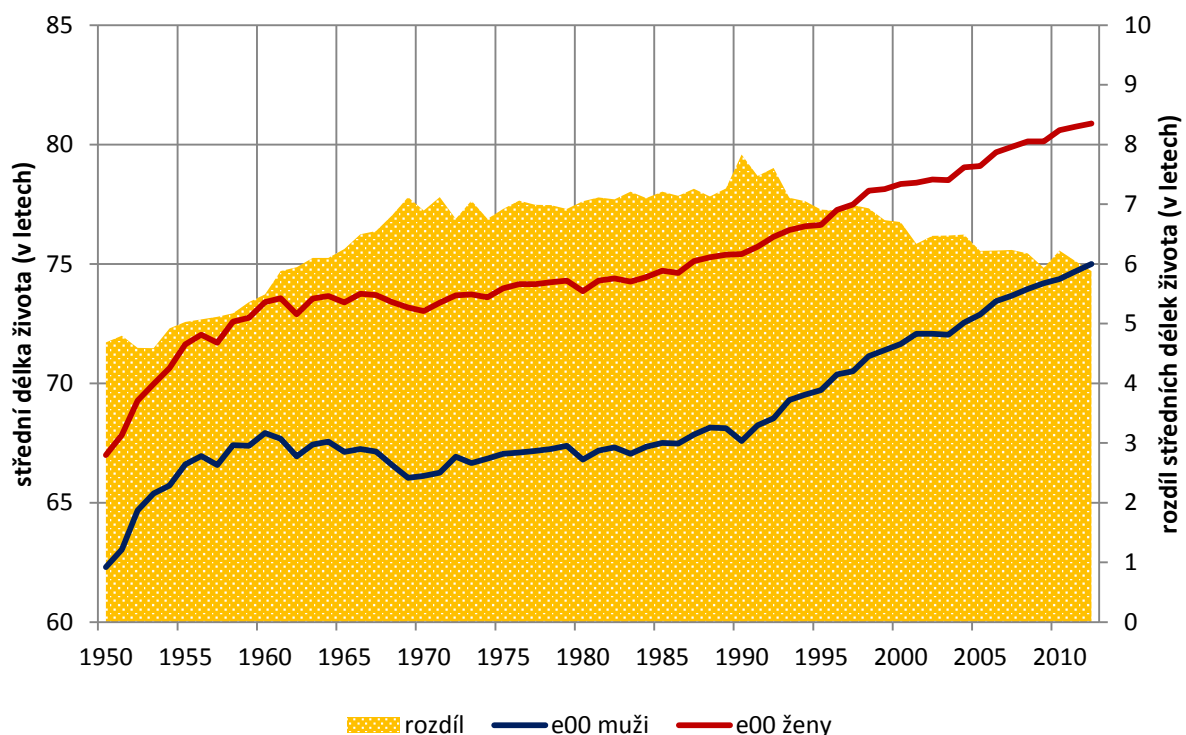
l_x, l_{x+n} – tabulkový počet dožívajících se ve věku x , resp. $x + n$ v letech A a B .

Tato metoda dekompozice se používá pro vyjádření změn v čase.

V tomto příspěvku nebudeme metodou dekompozice zjišťovat, jak se na rozdílu střední délky života mezi sledovanými roky podílely jednotlivé věkové skupiny, ale metodicky stejný postup výpočtu pro dekompozici použijeme na zjištění rozdílu středních délek života mezi ženami a muži ve vybraných letech.

Vývoj úmrtnosti mužů a žen v letech 1950–2012

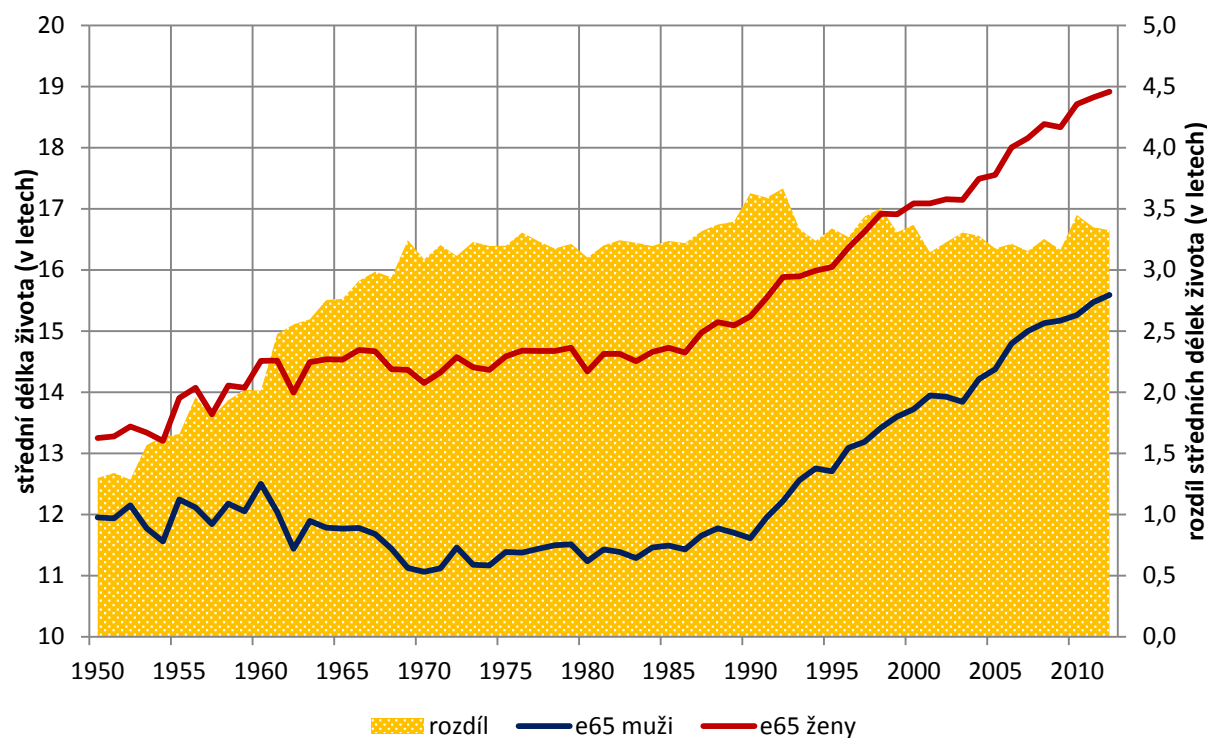
Česká republika patřila v dřívější době z pohledu úmrtnostních poměrů mezi demograficky vyspělé země. Ještě před druhou světovou válkou byla hodnota střední délky života mužů i žen s vyspělými zeměmi srovnatelná. V 50. letech 20. století v České republice výrazně poklesla kojenecká úmrtnost, především pak novorozenecká úmrtnost, což mělo za následek z více než 50% prodloužení střední délky života při narození. Vliv mělo především to, že v tomto období byl v České republice vybudován systém poraden pro matky a ony také začaly častěji rodit v lékařských zařízeních. Česká republika zavedla rovněž plošné očkování dětí. Došlo k významnému snížení úmrtnosti na nemoci infekční a parazitární. Do počátku šedesátých let patřila Česká republika mezi země s jednou z nejnižších kojeneckých úmrtností na světě.



Obrázek 27: Vývoj středních délek života pro 0 leté (e00) muže a ženy a rozdíl těchto délek života v letech 1950-2012

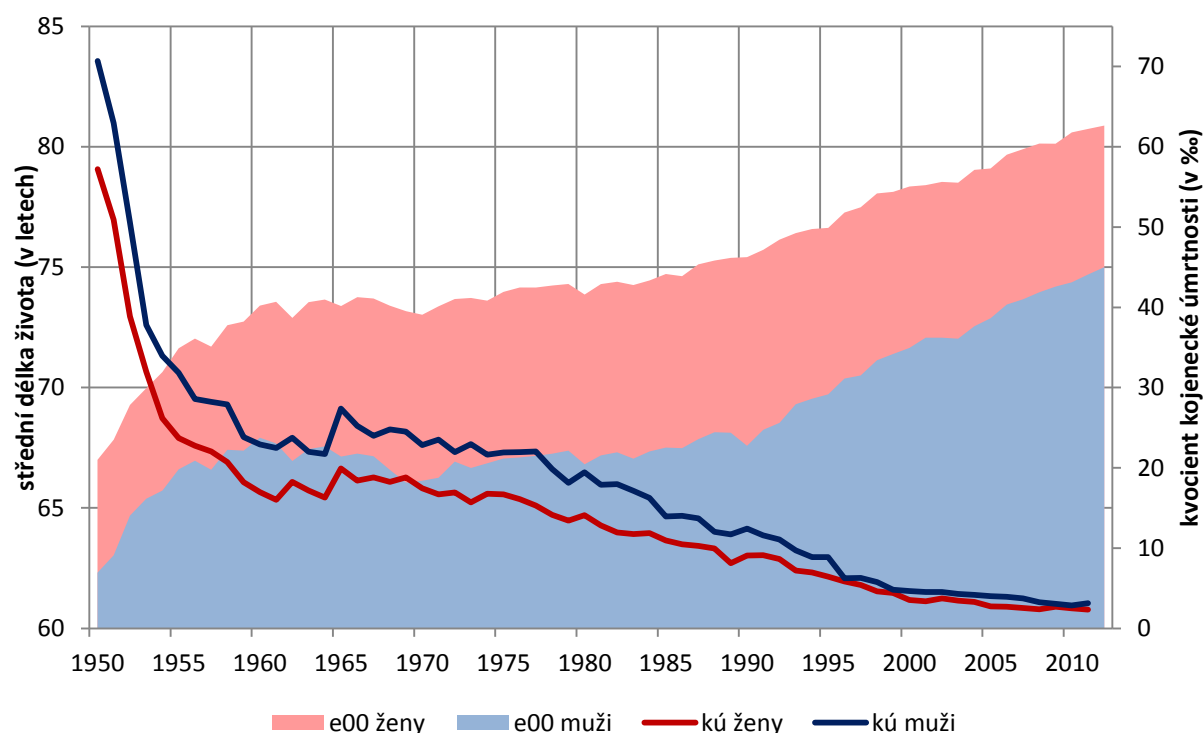
Zdroj: data ČSÚ, vlastní výpočty

Zhruba od poloviny šedesátých let dvacátého století do až do konce osmdesátých let se však úmrtnostní poměry ve srovnání s vyspělými státy zhoršily a Česká republika začala ve vývoji střední délky života zaostávat. V tomto období se pak politicky i demografickým vývojem přiřadila mezi tzv. socialistické východoevropské státy. Z pohledu úmrtnosti je patrná mužská nadúmrtnost ve středním a vyšším věku, také je zde horší úmrtnost na nemoci oběhové soustavy. Následkem těchto změn ve vývoji úmrtnosti dochází dokonce v určitých obdobích ke zkracování střední délky života u mužů a ke stagnaci délky života u žen. Zvětšují se rozdíly v úmrtnosti žen a mužů z pohledu střední délky života při narození (viz obrázek 1) i pro ostatní věky (viz obrázek 2). Ovšem pro vyšší věky je obecně platné, že je zde menší rozdíl mezi střední délkou života mužů a žen.



Obrázek 28: Vývoj středních délek života pro 65 leté (e65) muže a ženy a rozdíl těchto délek života v letech 1950-2012

Zdroj: data ČSÚ, vlastní výpočty



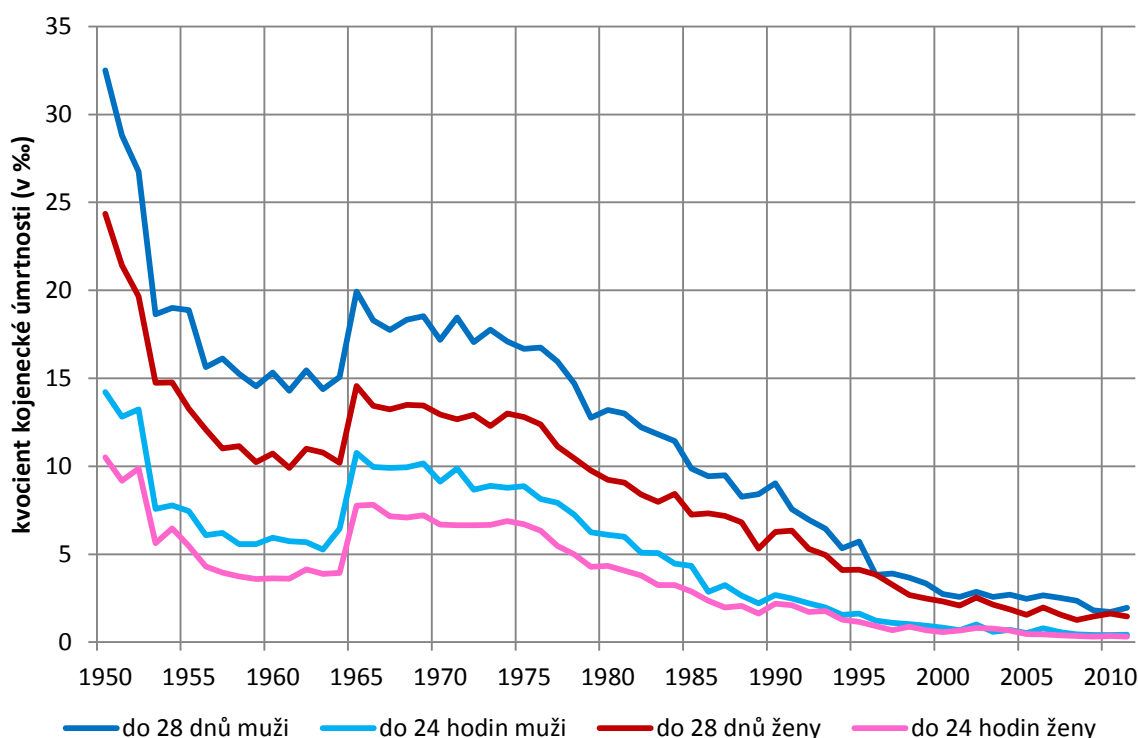
Obrázek 29: Porovnání středních délek života pro 0 leté (e00) muže a ženy kvocientu kojenecké úmrtnosti pro muže a ženy v letech 1950-2012

Zdroj: data ČSÚ, vlastní výpočty

Také lze pozorovat růst kojenecké úmrtnosti v roce 1965 (viz obrázek 3). Ten byl však způsoben změnou definice živě a mrtvě narozeného dítěte a potratu. Důvody proč dochází k zaostávání délky života mužů i žen za zeměmi západní Evropy v tomto období lze vysledovat v souvislosti s nedostatkem investic do zdravotnictví a zdraví obyvatelstva ve srovnání s vyspělými státy. Investice do vývoje nových léků,

prevence, nové operační postupy a metody, rozvoj technologií, zdravý životní styl byly v západoevropských státech výraznější a to se projevilo rychlejším zlepšováním úmrtnostních poměrů. Protože v České republice v období socialismu nebylo věnováno tolik prostředků do této oblasti, nedošlo k tak výrazné propagaci zdravého životního stylu a důrazu na zdravé životní a pracovní prostředí. Docházelo zde ke zhoršování, či zaostávání, z pohledu úmrtnosti za vyspělými státy.

Od konce sedmdesátých let dvacátého století dochází ke snižování kojenecké úmrtnosti v České republice. Do České republiky jsou dovezeny nové technologie, které jsou zaměřeny na novorozence a riziková těhotenství. Péče o matky a děti je na světové úrovni a kojenecká úmrtnost se postupně snižuje na hodnoty velmi příznivé. Koncem osmdesátých let dvacátého století se tyto změny v kojenecké úmrtnosti začínají výrazněji projevovat v růstu střední délky života při narození. I když trend vývoje úmrtnosti a střední délky života mužů i žen je příznivý, zaostáváme mírně za západoevropskými zeměmi. Z obrázku 3 je patrné, že po celé sledované období je kvocient kojenecké úmrtnosti chlapců vyšší než tento kvocient pro děvčata. To samé platí, i když se podíváme na vývoj úmrtnosti chlapců a děvčat do 28 dnů a 24 hodin po porodu (viz obrázek 4). Už tedy od narození můžeme pozorovat vyšší úmrtnost mužů oproti ženám.

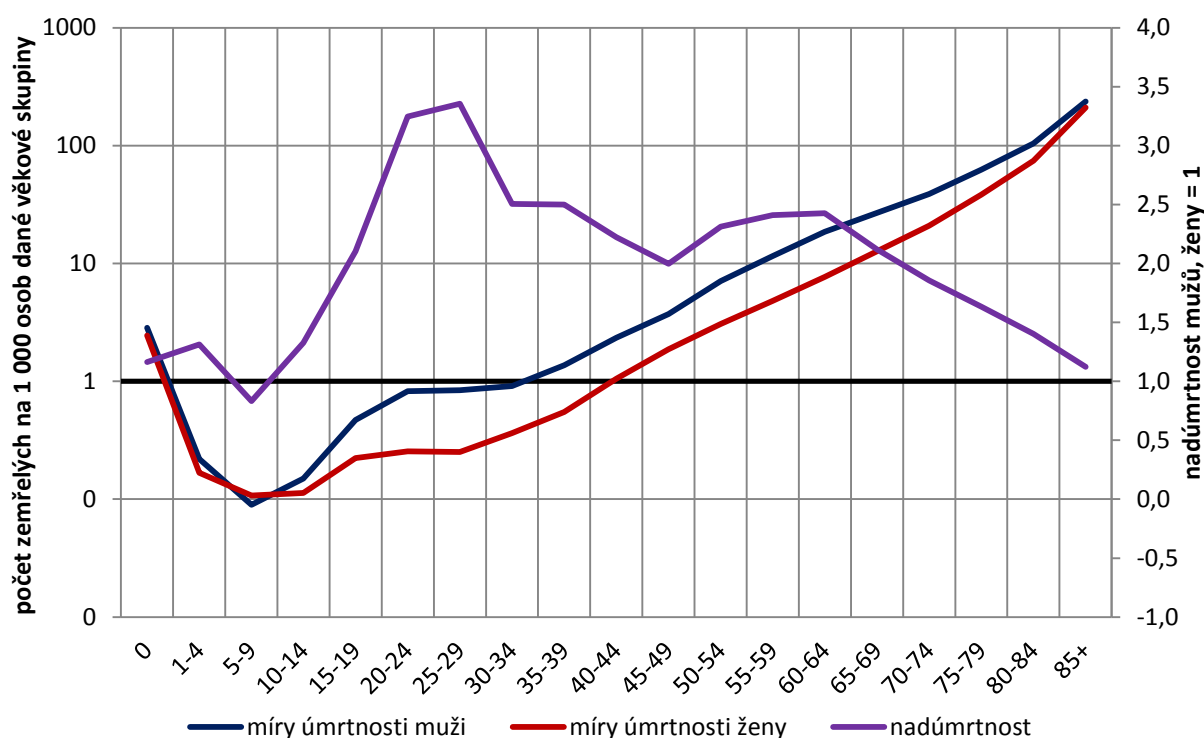


Obrázek 30: Porovnání úmrtnosti mužů a žen do 28 dnů a do 24 hodin v letech 1950-2012

Zdroj: data ČSÚ, vlastní výpočty

Změny, které později nastaly v úmrtnostních poměrech, se dávají do souvislosti s tzv. kardiovaskulární revolucí. V osmdesátých letech se i do České republiky postupně zavádějí nové postupy a metody v oblasti léčby a prevence nemocí oběhové soustavy. Rozvíjí se také způsoby rychlé zdravotnické pomoci z hlediska její dostupnosti, zlepšuje se informovanost obyvatelstva o možnostech prevence nemocí oběhové soustavy i jiných nemocí a lidé si více uvědomují nutnost investic do zdraví a prevence. Zvyšuje se důraz na zdravý životní styl, lepší pracovní a životní prostředí. Objevují se zde nové trendy v pohledu na zdraví obyvatelstva. Díky včasné diagnostice a novým možnostem léčby nemocí oběhové soustavy klesá úmrtnost v středním i vyšším věku na tuto příčinu. Z tohoto pohledu se pak snižuje rozdíl úmrtnosti mužů a žen. Klesá úmrtnost na příčiny smrti, které z dřívějšího pohledu byly neodvratitelné. Jak dochází k lepší dostupnosti nových léků, nejmodernějších léčebných metod, zlepšují se také úmrtnostní poměry mužů a žen v České republice. Snižuje se úmrtnost i ve vyšším věku, délka života mužů a žen roste a zmenšuje se mírně také rozdíl mezi střední délkou života žen a mužů. I přes tato zlepšení mírně zaostáváme z pohledu tohoto ukazatele za zeměmi západoevropskými, ke kterým se v šedesátých letech dvacátého století Česká republika z pohledu úmrtnostních poměrů řadila.

Lze však do budoucna předpokládat, že se střední délka života bude jak pro muže, tak pro ženy dále prodlužovat, ne však již z důvodu klesající kojenecké úmrtnosti (ta je již na velmi nízké a světově srovnatelné úrovni), ale z důvodu poklesu a zlepšení úmrtnosti ve středních a vyšších věcích. Rozdíly ve střední délce života žen a mužů se pak budou v důsledku těchto změn dále mírně snižovat.



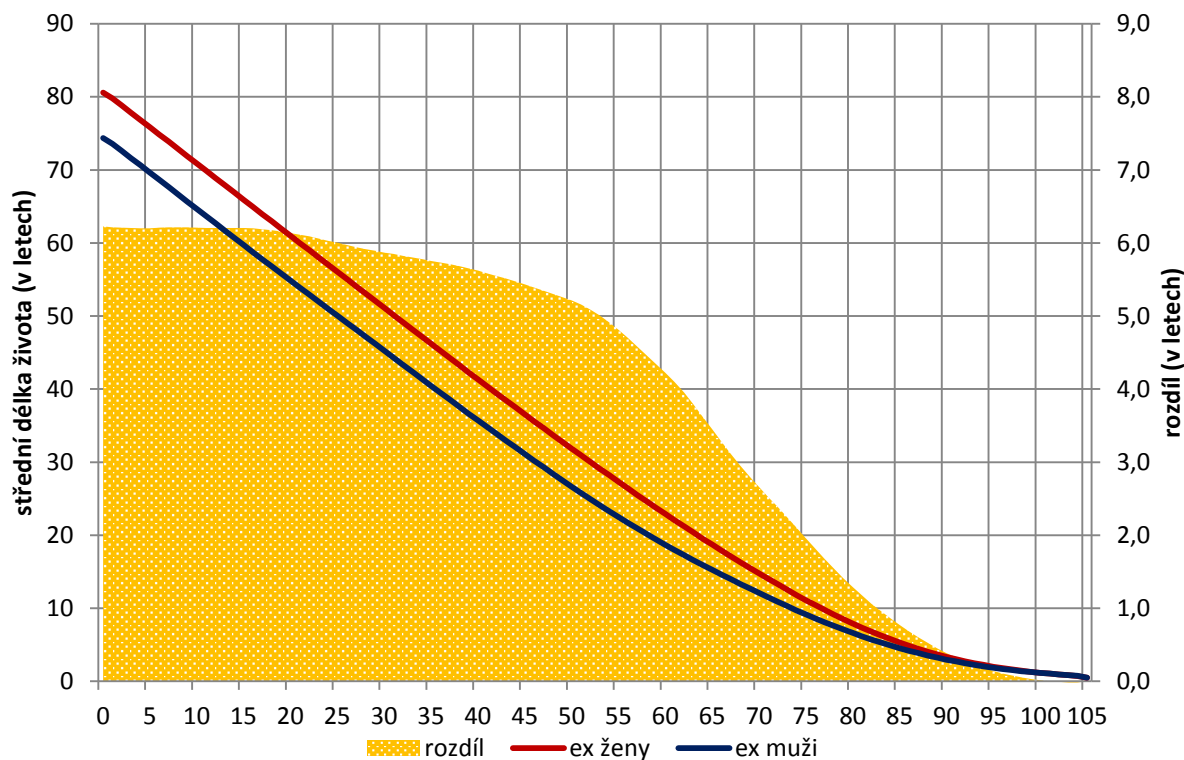
Obrázek 31: Míry úmrtnosti podle věku pro muže a ženy a mužská nadúmrtost v České republice v roce 2010

Zdroj: data ČSÚ, vlastní výpočty

Na obrázku 5 jsou zobrazeny specifické míry úmrtnosti pro muže a ženy za jednotlivé věkové skupiny v roce 2010. Rok 2010 byl zvolen na ukázkou, aby zde bylo možné zobrazit rozdíly mezi specifickou úmrtností mužů a žen. Graf je konstruován nikoli pro jednotlivé věky, ale pro víceleté věkové skupiny a to z důvodu výpočtu na více zastoupených skupinách. V tomto případě je patrné, že ve věkové skupině 5-9 let měly vyšší úmrtnost dívky. Ovšem nejedná se žádné pravidlo, konkrétně v tomto roce máme zaznamenáno více úmrtí dívek v daném věku, ale je zde také v porovnání s jinými věky podstatně méně pozorování. Dále je zde křivka zobrazující index nadúmrtosti, ze kterého je patrné kolikrát je úmrtnost mužů vyšší než úmrtnost žen pro jednotlivé věkové skupiny.

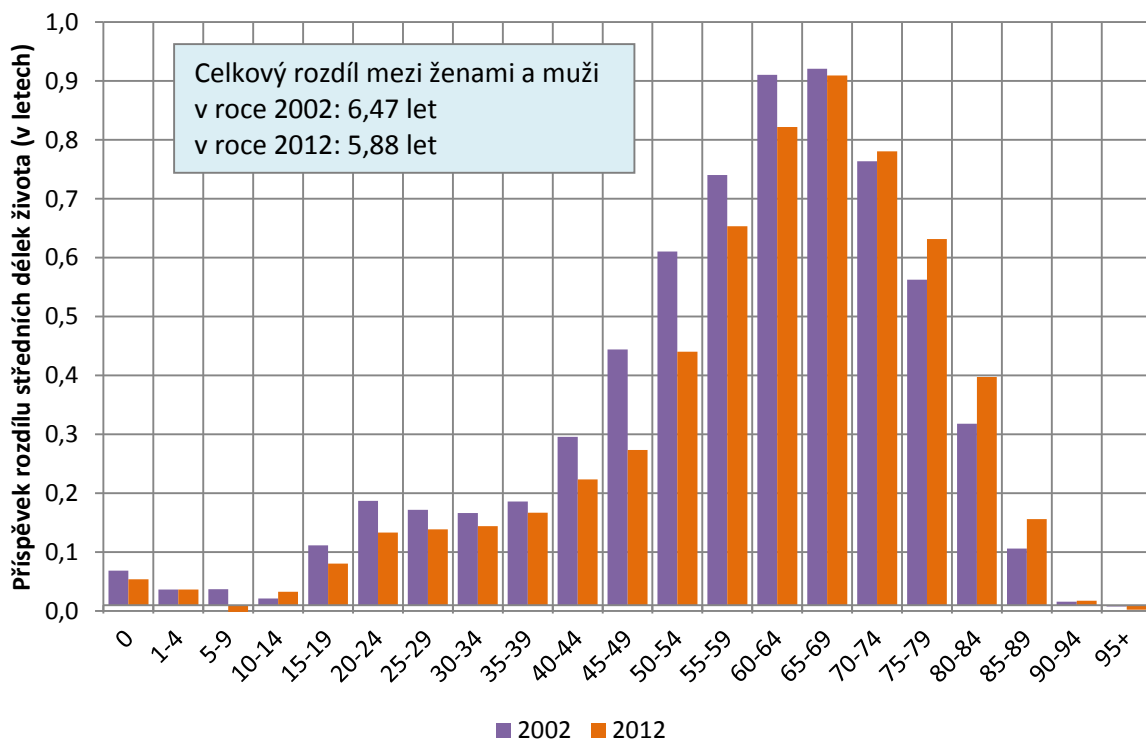
Dále je uvedeno srovnání středních délek života pro všechny věky v jednom roce, tím je opět rok 2010. Z obrázku 6 jasně vyplývá, že s přibývajícím rokem se rozdíly mezi střední délkou ženy v daném věku a muže ve stejném věku snižují. Například ve věku 84 let už mají žena a muž rozdíl ve střední délce života méně než rok.

V obrázcích 7 a 8 je zobrazen výsledek dekompozice rozdílu středních délek života mužů a žen. Výsledky byly zobrazeny na údajích z let 2002 a 2012, aby byl vidět posun za poslední dekádu. Z obrázku 7 je zřejmé, že v roce 2002 se na rozdíl ve střední délce života mezi ženami a muži podílela např. věková skupina 20-24 zhruba 0,2 roky. Do roku 2012 došlo k postupnému přeskupení vlivu jednotlivých skupin do vyšších věků.



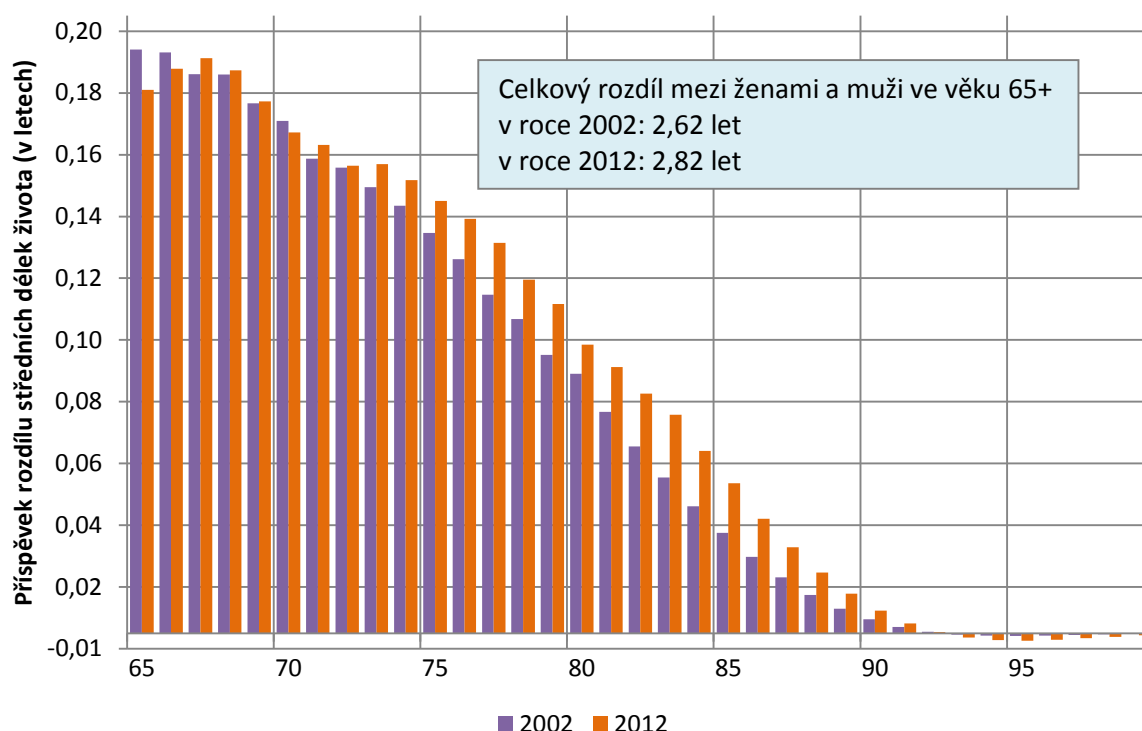
Obrázek 32: Střední délka života pro jednotlivé věky (ex) mužů a žen v České republice v roce 2010

Zdroj: data ČSÚ, vlastní výpočty



Obrázek 33: Příspěvky jednotlivých věkových skupin k rozdílu střední délky života při narození mezi ženami a muži v roce 2002 a 2012

Zdroj: data ČSÚ, vlastní výpočty



Obrázek 34: Příspěvky jednotlivých věkových skupin k rozdílu střední délky života 65 letých mezi ženami a muži v roce 2002 a 2012

Zdroj: data ČSÚ, vlastní výpočty

Závěr

Ve všech vyspělých zemích je z pohledu srovnání úmrtnosti mužů a žen patrná tzv. mužská nadúmrtost. I když se rodí o něco více chlapečků než dívek, ve starším věku převládají ženy. Je to způsobeno tím, že ve všech věcích mají muži vyšší intenzitu úmrtnosti než ženy. Dokonce již v kojeneckém věku lze vysledovat nadúmrtost chlapců. Mužská nadúmrtost se projevuje ve svém důsledku ve střední délce života při narození a její hodnotě. Na počátku 20. let minulého století byl rozdíl mezi střední délkou života mužů a žen při narození okolo 2,5 let a postupně tento rozdíl narůstal. Na počátku 60. let 20. století byl rozdíl mezi muži a ženami 5,5 roku a na počátku 90. let dosáhl téměř 8 let. Teprve v současnosti se stabilizuje na hodnotě okolo 6 let. Nejvyšší rozdíl ve střední délce života mezi ženami a muži je při narození, tento rozdíl postupně s věkem klesá. Pro nejstarší věky je minimální či již není rozdíl mezi pohlavími. Tedy platí, že pokud se muži dožijí vysokého věku, není již rozdíl mezi úmrtností podle pohlaví.

Do budoucna lze předpokládat, že se střední délka života bude jak pro muže, tak pro ženy dále prodlužovat, ne však již z důvodu klesající kojenecké úmrtnosti (ta je již na velmi nízké a světově srovnatelné úrovni), ale z důvodu poklesu a zlepšení úmrtnosti ve středních a vyšších věcích. Rozdíly ve střední délce života žen a mužů se pak budou v důsledku těchto změn dále mírně snižovat.

Literatura

Český statistický úřad. Dostupný z WWW: <<http://czso.cz/>>

FIALA, T.: *Výpočty aktuárské demografie v tabulkovém procesoru*. 1. vyd. Praha: Oeconomica, 2005. 177 s. ISBN 80-245-0821-4.

LANGHAMROVÁ, J.: *Úmrtnost v České republice v letech 1920-2011*. Diplomová práce, 2013.

MCCARTNEY, G., MAHMOOD, L., LEYLAND, A., BATTY, D., HUNT, K.: *Contribution of smoking-related and alcohol-related deaths to the gender gap in mortality: evidence from 30 European countries*, Tobacco Control, 2011, (Tob Control 2011;20:166-168 doi:10.1136/tc.2010.037929)

MEDICÍNA, Odborné fórum lékařů a farmaceutů.: *ICHS: rozdíly vmortalitě mužů a žen existují, ale jejich příčinou nejsou estrogeny*, MEDICÍNA 11, Roč. VIII, s. 7. Dostupný z WWW: < <http://www.zdrava-rodina.cz/med/med1101/med1113.html> >

PRESSAT, R.: *L'analyse démographique*, PUF 1961, 1969, 1983.

PRESSAT, R.: *Demographic analysis: methods, results, applications*, Chicago, 1972.

ROUBÍČEK, V.: *Úvod do demografie*. 1. vyd. Praha: CODEX Bohemia, 1997, 352 s. ISBN 80-85963-43-4.

STÁTNÍ ZDRAVOTNÍ ÚSTAV. *Hodnocení zdravotního stavu, Vybrané ukazatele demografické a zdravotní statistiky, Odborná zpráva za rok 2005*. Státní zdravotní ústav, Praha, srpen 2006.

TALACKO, J.: *Dynamická pozorování ve statistice úmrtnosti*. Praha: Ústřední statistický úřad, 1941.

JEL Classification: J110

Summary

Excess mortality of men in the Czech Republic

In terms of mortality compared to men and women we can see in all developed countries the male excess mortality. Although are born slightly more boys than girls in the older age dominate women. It is due to the fact that in all ages men have a higher intensity of mortality than women. Even in infant age can be observed excess mortality of boys. The male excess mortality is reflected ultimately in life expectancy at birth and its value. At the beginning of the 20th years of the last century was the difference between the life expectancy of men and women at birth of 2.5 years, and this difference gradually increased. At the beginning of the 60th years was a difference between men and women and 5.5 years at the beginning of the 90th years, reaching almost 8 years. Now stabilizes it at a value of about 6 years. The biggest difference in life expectancy between men and women at birth, this difference gradually decreases with age. For the oldest ages is minimal or is not the difference between the sexes. Therefore, if men live to old age, there is no difference between mortality rates by sex.

In the future, we can assume that life expectancy will be for both men and women continue to increase, not merely due to declining infant mortality (which is already very low and comparable to the world level), but because of the decline in mortality and improvements in secondary and higher ages. Differences in life expectancy between men and women will then be a consequence of these changes also slightly lower.

Key words: male excess mortality, life expectancy, infant mortality, quotient of infant mortality.

Metódy neparametrickej interpolácie dát

Tomáš Marcinko

xmart14@vse.cz

Doktorand oboru statistika

Školiteľ: doc. Ing. Dagmar Blatná, CSc. (blatna@vse.cz)

Abstrakt: Neparametrické regresné metódy, ktorých cieľom je dostatočne flexibilne popísať vzťah medzi veličinou nášho záujmu a jednou alebo viacerými sledovanými veličinami, sú užitočné hlavne v prípadoch, kedy je skúmaná závislosť príliš komplikovaná alebo nejednoznačná, resp. v situáciách, kedy nemáme k dispozícii dostatočne kvalitný parametrický model alebo vopred nepoznáme funkčný tvar skúmanej závislosti. Cieľom tohto príspevku je bližší popis významných neparametrických metód, a to Nadarayaovho-Watsonovho odhadu, lokálne polynomickej regresie, splajnového vyhladzovania a metódy penalizovaných splajnov.

Veľký dôraz pri použití týchto metód je potreba klásť na voľbu vyhladzovacieho parametra, ktorá má kľúčový vplyv na výsledný odhad regresnej funkcie. Preto si uvedieme dva hlavné prístupy, ako môžeme spoľahlivo určiť vyhladzovací parameter výhradne na základe dát. Odhady získané použitím týchto metód si z praktického hľadiska ilustrujeme na príklade závislosti logaritmu mzdy a veku jedinca. Ukážeme si, že väčšina uvažovaných metód s výnimkou Nadarayaovho-Watsonovho odhadu poskytuje v bežných situáciách dostatočne vhodné a flexibilné odhady regresnej funkcie. Na záver príspevku si zhrnieme silné stránky, ako aj potenciálne problémy týchto metód.

Kľúčová slova: neparametrická regresná analýza, lokálne polynomickej regresia, splajnové vyhladzovanie, penalizované splajny.

Úvod

Pri riešení reálnych problémov v mnohých oblastiach vedy sa často stretávame s potrebou popísať závislosť nejakej sledovanej náhodnej veličiny (vysvetľovaná premenná) na jednej alebo viacerých vopred daných veličinách (vysvetľujúce premenné). Nielen v ekonomickej praxi sa k tomuto účelu často využívajú hlavne lineárne modely, ktoré implicitne predpokladajú, že vplyv vysvetľujúcich premenných alebo ich transformácií na vysvetľovanú premennú je lineárny.

V reálnych aplikáciách však môže byť vzťah medzi skúmanými premennými príliš komplikovaný na to, aby mohol byť dostatočne kvalitne popísaný lineárnym modelom. Ak navyše nepoznáme ani typ funkčnej závislosti medzi skúmanými premennými, existuje možnosť využiť k popisu závislosti metódy neparametrickej regresie.

Odhad podmienenej strednej hodnoty vysvetľovanej premennej, ktorý je získaný na základe metód neparametrickej regresie, nezávisí na predpoklade funkčného tvaru závislosti, čím máme k dispozícii vyššiu mieru flexibility pri modelovaní vzťahu medzi vysvetľovanou premennou a vysvetľujúcimi premennými. Cenou za túto flexibilitu je hlavne vyššia výpočtová náročnosť, ako aj možné problémy neparametrického odhadu v prípade vyššieho počtu vysvetľovaných premenných. Neparametrické regresné metódy sú totiž často založené na myšlienke lokálneho priemerovania, ktoré môže z dôvodu nedostatku alebo riedkosti dát dávať v prípadoch viacerých vysvetľujúcich premenných neuspokojivé výsledky, prípadne pre niektoré kombinácie hodnôt vysvetľujúcich premenných nemusí byť odhad podmienenej strednej hodnoty vysvetľovanej premennej vôbec definovaný. V tomto článku sa však zameriame iba na prípad jedinej vysvetľujúcej premennej.

Metódy neparametrickej regresie

Predpokladajme, že máme k dispozícii realizáciu náhodného výberu $(x_1, y_1), \dots, (x_n, y_n)$, kde n udáva rozsah výberu. Neparametrická regresia všeobecne predpokladá, že podmienenú strednú hodnotu vysvetľovanej premennej Y za podmienky, že vysvetľujúca premenná X je rovná hodnote x , je možné vyjadriť pomocou regresného vzťahu

$$E(Y | x) = m(x),$$

kde m je neznáma hladká funkcia a x je príslušná hodnota vysvetľujúcej premennej. V tejto kapitole si uvedieme vybrané neparametrické metódy, ktoré umožňujú funkciu m odhadnúť bez apriórnej znalosti jej funkčného tvaru.

Jadrová regresia (Kernel regression)

Ako sme uviedli v úvode, viaceré metódy neparametrickej regresie sú založené na myšlienke lokálneho priemerovania. Táto idea vychádza z jednoduchej úvahy, že pozorovania (x_i, y_i) , u ktorých je hodnota x_i blízka hodnote x , by mali obsahovať viac informácie o hodnote $m(x)$ než vzdialenejšie pozorovania. Odhad regresnej funkcie v bode x by teda mal odpovedať váženému priemeru hodnôt y_i , pre ktoré je x_i v nejakom okolí bodu x . Samotné lokálne priemerovanie môžeme formálne vyjadriť ako

$$\hat{m}(x) = \frac{1}{n} \sum_{i=1}^n W_{ni}(x) y_i,$$

kde $W_{ni}(x)$ je váhová funkcia, pričom často platí, že

$$\frac{1}{n} \sum_{i=1}^n W_{ni}(x) = 1.$$

Jednou z prvých a najjednoduchších metód neparametrickej regresie je nepochybne *jadrová regresia*, ktorá využíva vyjadrenie váhovej funkcie $W_{ni}(x)$ pomocou tzv. *jadrovej funkcie* K . Táto jadrová funkcia je zväčša definovaná ako spojitá, obmedzená a symetrická reálna funkcia, pre ktorú platí

$$\int_{-\infty}^{\infty} K(u) du = 1.$$

Nielen pre účely jadrovej regresie bolo navrhnutých niekoľko jadrových funkcií, ukázalo sa však (viď napríklad Härdle (1990)), že výber jadrovej funkcie má iba malý vplyv na výsledný odhad regresnej funkcie. Z dôvodu niektorých optimálnych vlastností vzhľadom k tzv. asymptotickej strednej integrovanej štvorcovej chybe sa odporúča používať napríklad *Epanečnikovo jadro*

$$K(u) = \frac{3}{4} (1 - u^2) I_{[|u| < 1]}.$$

V prípade jadrovej regresie má váhová funkcia $W_{ni}(x)$ najčastejšie tvar

$$W_{ni}(x) = \frac{h_n^{-1} K[(x - x_i)/h_n]}{n^{-1} \sum_{i=1}^n h_n^{-1} K[(x - x_i)/h_n]},$$

kde h_n je tzv. vyhladzovací parameter (šírka okienka), u ktorého často predpokladáme, že pre všetky n platí $h_n = h$. Po dosadení tejto váhovej funkcie dostávame tzv. *Nadarayaov-Watsonov odhad* (viď Nadaraya (1964) a Watson (1964)) v tvare

$$\hat{m}_h(x) = \frac{\sum_{i=1}^n K[(x - x_i)/h] y_i}{\sum_{i=1}^n K[(x - x_i)/h]}.$$

Vplyv vyhladzovacieho parametra h je zrejmý: ak sa h bude blížiť nule, tak sa odhady regresnej funkcie v bode x_i budú blížiť hodnotám y_i , pričom rozptyl tohto odhadu bude vysoký. Naopak v prípade, že sa h bude blížiť nekonečnu, odhad regresnej funkcie sa pre ľubovoľné x bude blížiť k priemeru z hodnôt y_i , pričom tento odhad bude mať vysoké skreslenie. Voľba vyhladzovacieho parametra je teda založená na kompromise medzi rozptylom a skreslením odhadu regresnej funkcie.

Detaily o tejto metóde nájde záujemca napríklad v monografii Härdle (1990).

Lokálne polynomická regresia (Local polynomial regression)

Myšlienka lokálne polynomickej regresie je založená na tom, že hodnoty odhadovanej regresnej funkcie by mali približne odpovedať odhadu polynomickej regresnej funkcie v nejakom okolí bodu x . Tieto odhady sú všeobecne založené na minimalizácii vzťahu

$$\sum_{i=1}^n [y_i - \beta_0 - \beta_1(x - x_i) - \dots - \beta_p(x - x_i)^p]^2 K[(x - x_i)/h]$$

vzhľadom k vektoru parametrov $(\beta_0, \beta_1, \dots, \beta_p)'$, kde p vyjadruje zvolený rad polynómu, K je opäť zvolená jadrová funkcia a h je vyhladzovací parameter. Samotný odhad regresnej funkcie v bode x získame ako odhad neznámeho parametra β_0 , tzn. platí

$$\hat{m}_{p,h}(x) = \hat{\beta}_0(x) = (1, 0, \dots, 0)(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{y},$$

pričom využívame maticový zápis

$$\mathbf{X} = \begin{bmatrix} 1 & x - x_1 & \dots & (x - x_1)^p \\ 1 & x - x_2 & \dots & (x - x_2)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x - x_n & \dots & (x - x_n)^p \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

$$\mathbf{W} = \begin{bmatrix} K[(x - x_1)/h] & 0 & \dots & 0 \\ 0 & K[(x - x_2)/h] & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & K[(x - x_n)/h] \end{bmatrix}.$$

Najčastejšie využívané rady lokálne polynomickej regresnej funkcie sú $p = 1$ (*lokálne lineárna regresia*) alebo $p = 3$ (*lokálne kubická regresia*), pričom vyšší rad polynómu než 1 sa odporúča v prípade vysokej volatility dát. Uspokojivé výsledky môže poskytnúť i rad polynómu $p = 2$ (*lokálne kvadratická regresia*), príslušné odhady regresnej funkcie však majú horšie asymptotické vlastnosti, hlavne v blízkosti hraníc výberového priestoru hodnôt vysvetľujúcej premennej. Je vhodné poznamenať, že vyššie uvedený Nadarayaov-Watsonov odhad je špeciálnym prípadom lokálne polynomickej regresie v prípade voľby $p = 0$.

Vplyv vyhladzovacieho parametra h je obdobný, ako tomu bolo v prípade jadrovej regresie: ak sa h bude blížiť nule, tak sa odhady regresnej funkcie v bode x_i budú blížiť hodnotám y_i , pričom rozptyl tohto odhadu sa bude zvyšovať. Naopak v prípade, že sa h bude blížiť nekonečnu, odhad regresnej funkcie v ľubovoľnom bode x bude konvergovať k odhadu polynomickej regresnej funkcie radu p , ktorý získame pomocou klasickej metódy najmenších štvorcov.

Teoretické detaily k tejto metóde uvádza napríklad Simonoff (1996).

Splajnové vyhladzovanie (Spline smoothing)

Na rozdiel od prvých dvoch metód neparametrickej regresie, ktoré sme si uviedli, tzv. *splajnové vyhladzovanie* nevyužíva k odhadu regresnej funkcie $m(x)$ jadrovú funkciu. Podobne, ako je tomu v prípade klasickej regresie, uvažujeme minimalizáciu reziduálneho súčtu štvorcov, avšak za podmienky dostatočnej hladkosti funkcie $m(x)$. Splnenie tejto podmienky dosiahneme zahrnutím penalizačného člena do vzťahu pre reziduálny súčet štvorcov, pričom sa najčastejšie využíva penalizačný člen v tvare

$$\int_{-\infty}^{\infty} [m''(x)]^2 dx.$$

Zmyslom tohto penalizačného člena je, že s rastúcou krivosťou regresnej funkcie $m(x)$ rastie i hodnota $|m''(x)|$, čím rastie i hodnota daného integrálu. Odhad neparametrickej regresnej funkcie je teda založený na minimalizácii vzťahu

$$\sum_{i=1}^n [y_i - m(x_i)]^2 + h \int_{-\infty}^{\infty} [m''(x)]^2 dx,$$

kde h spĺňa funkciu vyhladzovacieho parametra. Vplyv vyhladzovacieho parametra h je opäť obdobný, ako tomu bolo v predošlých prípadoch: ak sa h bude blížiť nule, tak sa odhady regresnej funkcie v bode x_i budú blížiť hodnotám y_i , pričom rozptyl takéhoto odhadu bude vysoký. Naopak v prípade, že sa h bude blížiť nekonečnu, odhad regresnej funkcie bude v každom bode x konvergovať k odhadu lineárnej regresnej funkcie, ktorý získame použitím klasickej metódy najmenších štvorcov.

Získanie odhadu regresnej funkcie však v prípade splajnového vyhladzovania nie je tak priamočiary, ako tomu bolo napríklad u lokálne polynomickej regresie. Dá sa však ukázať, že ak uvažujeme triedu aspoň dvakrát diferencovateľných funkcií na uzavretom intervale $[x_{(1)}, x_{(n)}]$, tak vyššie uvedený vzťah pre reziduálny súčet štvorcov s penalizačným členom minimalizuje tzv. *kubický splajn*, ktorý tvoria kubické polynómy medzi susednými hodnotami $x_{(i)}$ a $x_{(i+1)}$, kde $x_{(i)}$ značí i -tú poradovú štatistiku. Navyše platí, že kubický splajn je lineárnou funkciou pozorovaní y_i , a teda existujú váhy $W_{hi}(x)$ také, že platí

$$\hat{m}_h(x) = \frac{1}{n} \sum_{i=1}^n W_{hi}(x) y_i.$$

K určeniu samotného odhadu kubického splajnu sa využívajú špeciálne metódy, napríklad Reinschov algoritmus (viď Reinsch (1967) alebo Green a Silverman (1994)). Detailný teoretický rozbor splajnového vyrovnávania uvádza napríklad Eubank (1999).

Penalizované splajny (Penalized spline regression)

Metóda penalizovaných splajnov, ktorú navrhli Ruppert, Wand a Carroll (2003), je založená na splajnovom modeli (radu p), ktorý má explicitný tvar

$$m(x) = \beta_0 + \beta_1 x + \dots + \beta_p x^p + \sum_{k=1}^K \beta_{pk} (x - \kappa_k)_+^p,$$

kde $\beta = (\beta_0, \beta_1, \dots, \beta_p, \beta_{p1}, \dots, \beta_{pK})'$ je vektor reálnych parametrov, $\kappa_1, \dots, \kappa_K$ sú tzv. uzly vyjadrujúce ľubovoľne umiestnené body vo výberovom priestore vysvetľujúcej premennej, K je počet uzlov a značenie $(x - \kappa_k)_+$ vyjadruje nezápornú časť výrazu v zátvorke, tzn. platí

$$(x - \kappa_k)_+ = \max\{0; x - \kappa_k\}.$$

Aj keď počet i umiestnenie uzlov $\kappa_1, \dots, \kappa_K$ môže byť ľubovoľné, čo umožňuje analytikovi veľký stupeň flexibility, v bežných prípadoch často postačuje automatická voľba uzlov. Autori navrhli počet uzlov určený ako $K = \min\{1/4 \text{ počtu unikátnych pozorovaní } x_i; 35\}$, pričom tieto uzly môžu byť rozmiestnené v bodoch rovnomerne rozložených empirických kvantilov získaných z hodnôt $x_1, \dots, x_n$.

Vzhľadom k tomu, že použitie vyššie uvedeného splajnového modelu bez akýchkoľvek dodatočných obmedzení môže viesť k výrazne variabilnej regresnej funkcii, sa za účelom dosiahnutia dostatočne hladkej funkcie požaduje splnenie podmienky v tvare

$$\sum_{k=1}^K \beta_{pk}^2 < C,$$

kde C je zvolená konštanta.

Odhad vektoru parametrov β v penalizovanom splajnovom modeli dostaneme riešením optimalizačnej úlohy, ktorú pri využití maticového zápisu

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & \dots & x_1^p & (x_1 - \kappa_1)_+^p & \dots & (x_1 - \kappa_K)_+^p \\ 1 & x_2 & \dots & x_2^p & (x_2 - \kappa_1)_+^p & \dots & (x_2 - \kappa_K)_+^p \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^p & (x_n - \kappa_1)_+^p & \dots & (x_n - \kappa_K)_+^p \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{D} = \begin{bmatrix} \mathbf{0}_{p+1, p+1} & \mathbf{0}_{p+1, K} \\ \mathbf{0}_{K, p+1} & \mathbf{I}_K \end{bmatrix},$$

môžeme zapísať ako

$$\min_{\boldsymbol{\beta}} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 \text{ za podmienky } \boldsymbol{\beta}'\mathbf{D}\boldsymbol{\beta} < C.$$

Riešením metódou Lagrangeových multiplikátorov dostaneme ekvivalentnú úlohu

$$\min_{\boldsymbol{\beta}} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + h^2 \boldsymbol{\beta}'\mathbf{D}\boldsymbol{\beta},$$

kde Lagrangeov multiplikátor $h \geq 0$ môže takisto plniť funkciu vyhladzovacieho parametra. Riešením minimalizačnej úlohy získavame explicitný odhad vektoru parametrov $\boldsymbol{\beta}$ v tvare

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X} + h^2\mathbf{D})^{-1}\mathbf{X}'\mathbf{y}.$$

Pre pozorované hodnoty $x_1, \dots, x_n$ teda dostávame odhad regresnej funkcie v tvare

$$\hat{\mathbf{m}} = \begin{bmatrix} \hat{m}_{p,h}(x_1) \\ \hat{m}_{p,h}(x_2) \\ \dots \\ \hat{m}_{p,h}(x_n) \end{bmatrix} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}'\mathbf{X} + h^2\mathbf{D})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{S}\mathbf{y},$$

z ktorého je zrejmé, že odhadnutý penalizovaný splajn je takisto lineárnou funkciou pozorovaní y_i .

V praxi sa častejšie než horná hranica penalizujúcej podmienky C volí hodnota vyhladzovacieho parametra h , pričom vplyv vyhladzovacieho parametra je opäť podobný, ako tomu bolo v prípade splajnového vyhladzovania: ak zvolíme parameter h rovný nule, dostaneme splajnový model radu p , ktorý v závislosti na počte a umiestnení uzlov môže byť takmer ľubovoľne variabilný. Naopak v prípade, že sa h bude blížiť nekonečnu, odhad regresnej funkcie bude v každom bode x konvergovať k odhadu polynomickej regresnej funkcie radu p , ktorú získame použitím klasickej metódy najmenších štvorcov. Rovnako, ako tomu bolo v prípade lokálnej polynomickej regresie, volíme rad polynómu p v závislosti na volatilitate dát.

Voľba vyhladzovacieho parametra

Pri využívaní metód neparametrickej regresie je kľúčová vhodná voľba vyhladzovacieho parametra. Jednou z možností je taká voľba vyhladzovacieho parametra, ktorú bude užívateľ neparametrických metód používať na základe svojich skúseností za vhodnú, je však možné využiť i rôzne algoritmy, pomocou ktorých je možné určiť hodnotu vyhladzovacieho parametra automaticky na základe príslušného dátového súboru.

Jedným z najčastejšie používaných prístupov je voľba vyhladzovacieho parametra na základe tzv. *křížového overovania* (cross-validation), ktoré je založené na minimalizácii funkcie vyhladzovacieho parametra h v tvare

$$CV(h) = \sum_{i=1}^n [y_i - \hat{m}_h^{(i)}(x_i)]^2,$$

kde $\hat{m}_h^{(i)}(x_i)$ značí odhad regresnej funkcie m v bode x_i po vynechaní i -tého pozorovania a pri voľbe vyhladzovacieho parametra h .

Ak je odhad regresnej funkcie m v bode x_i lineárnou funkciou pozorovaní vysvetľovanej premennej, tzn. platí

$$\hat{m}(x_i) = \sum_{j=1}^n s_{ij} y_j,$$

potom je možné odhad $\hat{m}_h^{(i)}(x_i)$ určiť ako

$$\hat{m}(x_i) = \frac{\sum_{j \neq i} s_{ij} y_j}{\sum_{j \neq i} s_{ij}}.$$

Vzhľadom k tomu, že voľba vyhladzovacieho parametra založená výhradne na minimalizácii funkcie $CV(h)$ môže viesť nižším hodnotám vyhladzovacieho parametra, bol navrhnutý alternatívny prístup, ktorý využíva penalizačný člen zahrnutý do výrazu pre strednú štvorcovú chybu, čo vedie k takej voľbe vyhladzovacieho parametra h , ktorá minimalizuje funkciu

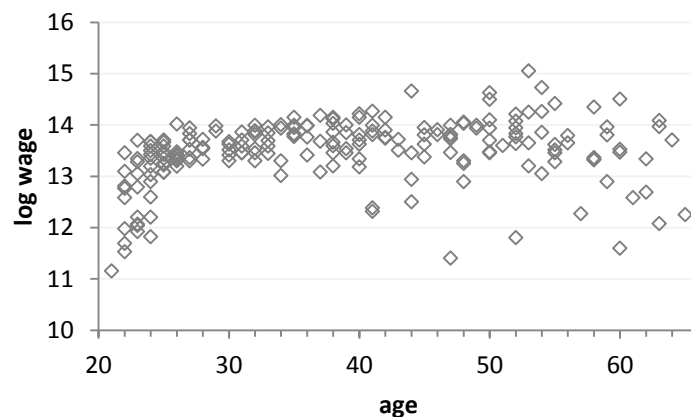
$$G(h) = \sum_{i=1}^n [y_i - \hat{m}_h(x_i)]^2 \Xi(n^{-1} W_{hi}(x_i)).$$

Použitím penalizovanej funkcie v tvare $\Xi_{GCV}(u) = (1 - u)^{-2}$ dostávame tzv. *zovšeobecnené krížové overovanie*, inými možnosťami sú tzv. Akaikeho kritérium založené na penalizačnej funkcii $\Xi_{AIC}(u) = \exp(2u)$, prípadne tzv. Riceovo T , ktoré je založené na funkcii $\Xi_T(u) = (1 - 2u)^{-1}$.

Teoretické odvodenie rôznych kritérií pre voľbu vhodnej hodnoty vyhladzovacieho parametra nájde záujemca v práci Härdle et al. (2004).

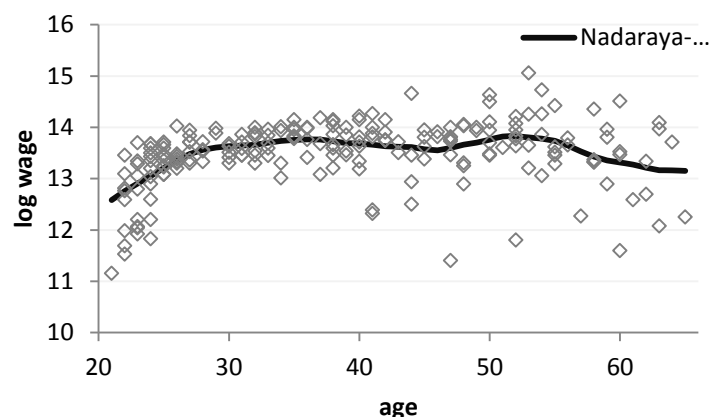
Aplikácia metód neparametrickej regresie

Metódy neparametrickej regresie, ktoré sme si uviedli, budeme ilustrovať na prierezových dátach, ktoré uvádzajú Pagan a Ullah (1999). Jedná sa o náhodný výber 205 pracujúcich mužov s dosiahnutým bežným stredoškolským vzdelaním – údaje pochádzajú z cenzu, ktoré bolo uskutočnené v Kanade v roku 1971. U týchto 205 osôb bola sledovaná závislosť medzi prirodzeným logaritmom mzdy a veku jedinca. Údaje sú zaznamenané na nasledovnom obrázku:



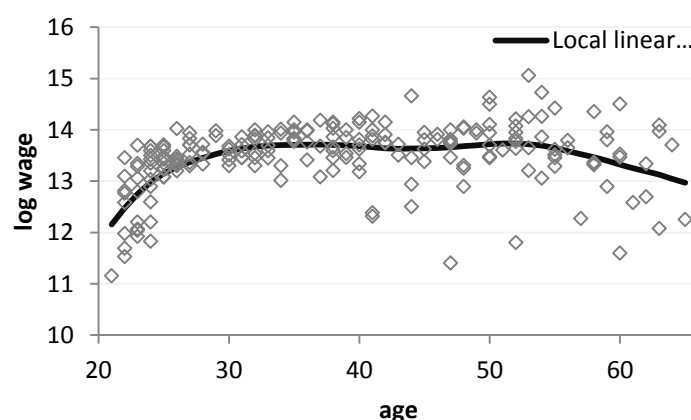
Obr. 1: Zobrazenie dát o veku a logaritme mzdy

Priebeh vzťahu medzi logaritmom mzdy a vekom sa najprv pokúsime popísať pomocou jadrovej regresie, resp. použitím lokálne polynomickej regresie s rôznymi radmi polynómu p . Vo všetkých prípadoch budeme používať Epanečnikovo jadro, pričom vyhladzovací parameter bol stanovený minimalizáciou funkcie $CV(h)$, prípadne funkcie $G(h)$ s využitím $\Xi_T(u)$. Oba prístupy viedli často k veľmi podobným výsledkom.

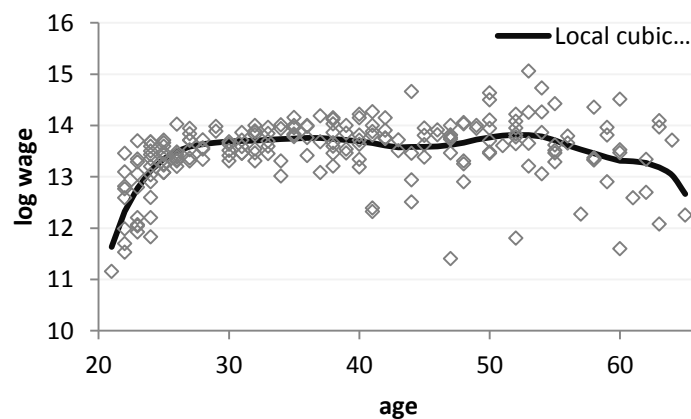


Obr. 2: Nadarayaov-Watsonov odhad regresnej funkcie ($h_{CV} = 3,4$)

Ako vidíme na obrázku 2, Nadarayaov-Watsonov odhad regresnej funkcie $m(x)$ síce dáva dostatočnú informáciu o priebehu závislosti medzi logaritmom mzdy a vekom jedinca, samotný odhad však nemožno považovať za plne žiaduci vzhľadom k podmienke dostatočnej hladkosti regresnej funkcie. Takisto si môžeme všimnúť, že lokálne skreslenie na ľavej hranici výberového priestoru hodnôt vysvetľujúcej premennej sa zdá byť pomerne výrazné. Toto lokálne skreslenie vzniká z dôvodu, že v okolí minimálnej hodnoty $x_{(1)}$, resp. maximálnej hodnoty $x_{(n)}$, sa nachádzajú pozorovania, na základe ktorých je odhad regresnej funkcie v týchto bodoch stanovený, len na pravej, resp. ľavej strane príslušného okolia, pričom príslušné váhy ostávajú symetrické.



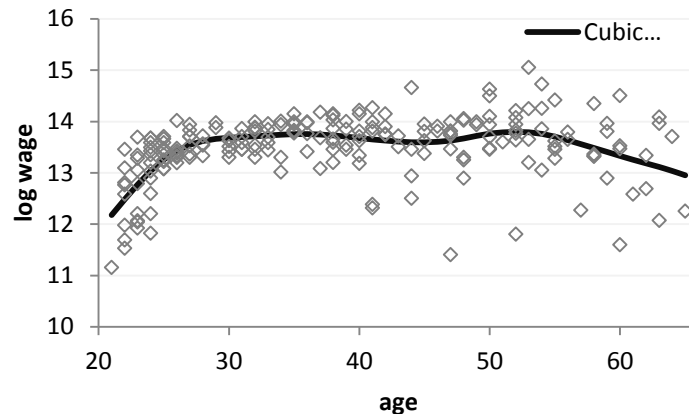
Obr. 3: Lokálne lineárny odhad regresnej funkcie ($h_{CV} = 7,3$)



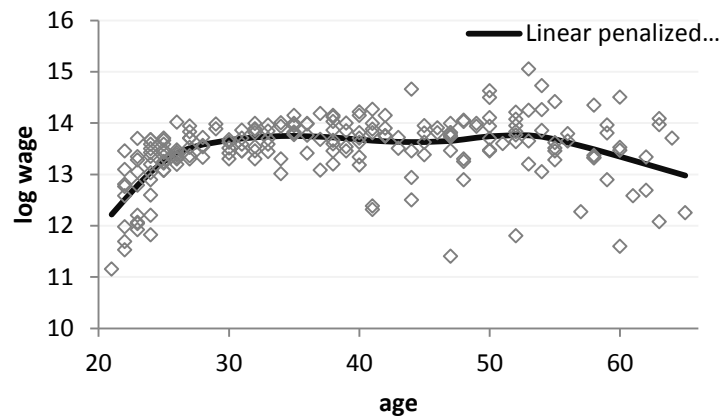
Obr. 4: Lokálne kubický odhad regresnej funkcie ($h_T = 8,7$)

Na obrázkoch 3 a 4 môžeme vidieť, že metóda lokálnej polynomickej regresie s nenulovým radom p vedie k vhodnejším odhadom regresnej funkcie $m(x)$. Z grafického posúdenia dát vychádza lepšie lokálne lineárny odhad regresnej funkcie, ktorý poskytuje dostatočne hladký odhad regresnej funkcie s rozumným kompromisom medzi rozptylom a skreslením odhadu $m(x)$. Vzhľadom k tomu, že závislosť medzi sledovanými veličinami nebude príliš komplikovaná, nie je opodstatnené použiť vyšší rad polynómu p než 1. Z obrázku 4 sa zdá, že lokálne kubická regresia vedie k až príliš variabilnému odhadu regresnej funkcie s pomerne výrazným skreslením na oboch hraniciach výberového priestoru hodnôt vysvetľujúcej premennej – platí totiž, že zvýšením radu polynómu p spravidla rastie i skreslenie odhadu $m(x)$ pri hraniciach výberového priestoru.

Na nasledovných obrázkoch si ukážeme, ako vyzerajú príslušné odhady regresnej funkcie získané kubickým alebo penalizovaným splajnom.

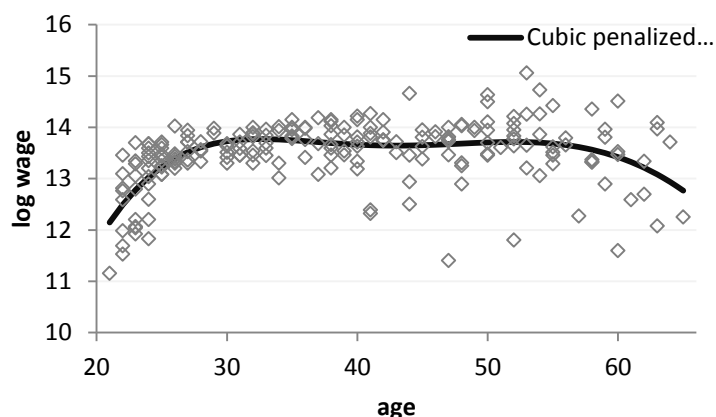


Obr. 5: Kubický splajn ($h_{CV} = 0,6$)



Obr. 6: Lineárny penalizovaný splajn ($h_{CV} = 12,8$)

Pri pohľade na obrázky 5 až 7 je zrejmé, že každý z príslušných odhadov regresnej funkcie možno považovať za dostatočne kvalitný odhad regresnej funkcie popisujúcej závislosť medzi logaritmom mzdy a vekom jedinca. Zjavná je taktiež výrazná podobnosť medzi kubickým splajnom a lineárnym penalizovaným splajnom (pripomeňme, že s rastúcou hodnotou vyhladzovacieho parametra h konvergujú kubický splajn i lineárny penalizovaný splajn k rovnakému odhadu lineárnej regresnej funkcie získanému metódou najmenších štvorcov). Vzhľadom k jednoduchosti odhadnutej regresnej funkcie, ako aj prirodzenému chovaniu tejto funkcie na pravej hranici výberového priestoru hodnôt vysvetľujúcej premennej, sa dá považovať za najpriateľnejší odhad kubickým penalizovaným splajnom.

Obr. 7: Kubický penalizovaný splajn ($h_{CV} = 15,2$)

Záver

V druhej polovici 20. storočia bolo navrhnutých niekoľko neparametrických metód regresnej analýzy, ktoré tvoria alternatívu pre parametrické lineárne alebo nelineárne modely, a to predovšetkým v prípadoch, kedy je závislosť medzi skúmanými veličinami komplikovaná či nejednoznačná. V tomto príspevku sme sa zamerali na štyri z nich: Nadarayaov-Watsonov odhad, lokálne polynomicnú regresiu, kubický splajn a penalizované splajny.

Na ilustračnom príklade sme si ukázali, že najstaršia z popísaných metód (Nadarayaov-Watsonov odhad) neposkytuje dostatočne hladký odhad regresnej funkcie, a tak túto metódu jadrovej regresie nemôžeme odporučiť. Naopak lokálne lineárny odhad, kubický splajn i penalizované splajny dávajú z praktického pohľadu veľmi podobné a dostatočne presné odhady regresnej funkcie. Aj keď z osobného hľadiska považujem za najužitočnejšiu a najflexibilnejšiu metódu penalizovaných splajnov, voľba akejkoľvek z týchto neparametrických metód sa javí ako vhodná a konkrétny výber závisí skôr na preferenciách analytika či na prístupnosti výpočtových algoritmov, ktoré žiaľ často nie sú zahrnuté v štandardných štatistických paketoch.

Medzi silné stránky neparametrických regresných metód patrí hlavne:

- možnosť získania prvotnej informácie o tvare závislosti medzi sledovanými veličinami,
- interpolácia dát bez apriórnej znalosti funkčného tvaru závislosti medzi veličinami,
- odhad regresnej funkcie, ak z interpretačného hľadiska nie je potrebný parametrický model,
- možnosť dodatočnej konštrukcie konfidenčných a predikčných intervalov pre $m(x)$,
- možnosť verifikácie vhodnosti zvoleného lineárneho alebo nelineárneho modelu.

Vzhľadom k tomu, že metódy neparametrickej regresie sú s ohľadom na výkonnosť výpočtovej techniky bezproblémovo prístupné, je možné odporučiť použitie týchto metód za účelom verifikácie zvoleného parametrického modelu vždy v prvej fáze regresnej analýzy.

Pre úplnosť však musíme upozorniť i na možné problémy, ktoré sú s neparametrickou regresiou úzko spojené. Väčšinu nedostatkov je však možné vhodným spôsobom obmedziť alebo eliminovať.

- subjektívnosť voľby vyhladzovacieho parametra – nesprávna voľba vyhladzovacieho parametra môže viesť k nevhodnému odhadu regresnej funkcie,
- popísané metódy sú nevhodné pre extrapoláciu dát – za účelom extrapolácie dát je vhodnejšie využiť verifikovaný lineárny alebo nelineárny model,
- lokálne skreslenie na krajoch výberového priestoru hodnôt vysvetľujúcej premennej – možné obmedziť zvolením vhodných váh pre jednotlivé pozorovania, viď napríklad Härdle (1990),
- problémy s nedostatkom alebo riedkosťou dát v prípadoch viacerých vysvetľujúcich premenných – možné využiť semiparametrické modely, viď napríklad Härdle et al. (2004).

Literatura

- EUBANK, R. L. *Nonparametric Regression and Spline Smoothing*. New York: Marcel Dekker, 1999.
- GREEN, P. J., SILVERMAN, B. W. *Nonparametric Regression and Generalized Linear Models: A roughness penalty approach*. London: Chapman & Hall, 1994.
- HÄRDLE, W. *Applied Nonparametric Regression*. Cambridge: Cambridge University Press, 1990.
- HÄRDLE, W., MÜLLER, M., SPERLICH, S., WERWATZ, A. *Nonparametric and Semiparametric Models*. Berlin: Springer, 2004.
- NADARAYA, E. A. On estimating regression. *Theory of probability and its applications*. 1964, roč. 9, č. 1, s. 141-142.
- PAGAN, A., ULLAH, A. *Nonparametric Econometrics*. New York: Cambridge University Press, 1999.
- RUPPERT, D., WAND, M. P., CARROLL, R. J. *Semiparametric Regression*. New York: Cambridge University Press, 2003.
- REINSCH, H. Smoothing by spline functions. *Numerische Mathematik*. 1967, roč. 10, č. 3, s. 177-183.
- SIMONOFF, J. S. *Smoothing Methods in Statistics*. New York: Springer, 1996.
- WATSON, G. S. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*. 1964, roč. 26, č. 4, s. 359-372.

JEL Classification: C14.

Summary

Nonparametric Smoothing Methods

Nonparametric smoothing methods, which are designed to model a general relationship between a response variable and one or more explanatory variables in a sufficiently flexible way, are useful especially in those cases, where the given relationship is too complicated or ambiguous, or in situations, when we do not have at disposal a sufficiently good parametric model and / or we do not have any prior knowledge about a functional form of the given relationship. The aim of this article was to give a comprehensive description of the most important nonparametric methods, namely the Nadaraya-Watson estimator, local polynomial regression, spline smoothing and penalized splines.

When using these methods, the choice of a smoothing parameter has a crucial impact on the resulting estimate of the regression function. Therefore, we provided two main approaches, how to determine a reliable data-driven smoothing parameter. Estimates obtained by these nonparametric methods were illustrated and compared from a practical standpoint on a relationship between age of an individual and a logarithm of his wage. It was shown that most of the described methods, except for Nadaraya-Watson estimator, generate sufficiently desirable and flexible estimates of the regression function in ordinary situations. In the end of the article the main strengths, as well as potential drawbacks of these methods were shortly discussed.

Key words: nonparametric regression analysis, local polynomial regression, spline smoothing, penalized splines.

Statistické analýzy regionálních diferencí v oblasti udržitelného rozvoje a kvality života

Ludmila Petkovová

Ludmila.petkovova@vse.cz

Doktorand oboru statistika

Školitel: doc. Ing. Jakub Fischer, Ph.D., (jakub.fischer@vse.cz)

Abstrakt: Česká republika prochází v posledních 30 letech významnými ekonomickými i sociálními změnami, které se různí ve svých dopadech v jednotlivých regionech. Otázkou je, zdali tyto změny povedou k sbližování jednotlivých regionů ČR či naopak k prohlubování meziregionálních rozdílů.

Príspevek se zabývá možnostmi hodnocení regionálních diferencí v oblastech, které jsou vystihovány kvantitativně těžko zachytitelnými jevy s nutností využití většího množství ukazatelů. Indikátorové sady prezentující desítky až stovky údajů jsou v těchto ohledech nepřehledné a omezující a jejich použitím tak dochází ke ztrátě jednoduchosti a výstižnosti výsledku. Trendem posledních let je proto ústup od prezentace takto náročně pojatých výsledků a hledání nových cest, například v podobě kompozitních indikátorů. Hlavní výhoda využití kompozitních indikátorů spočívá v možnosti shrnutí komplexních jevů umožňující následně provést snadné porovnání regionů.

Príspevek je zaměřen především na metody sledování regionálních diferencí pomocí agregovaného indikátoru zahrnujícího značné množství nesourodých ukazatelů v podmínkách České republiky. Metody jsou ukázány na datech pro udržitelný rozvoj, velmi obdobné je však i použití v případech hodnocení kvality života. Ta je i přes odlišnost své koncepce z kvantitativního hlediska velmi podobná právě udržitelnému rozvoji.

Klíčová slova: Regionální difference, Kompozitní indikátory, Regionální analýza

Úvod

Sledování vývoje regionů, jejich sbližování či naopak vzájemné vzdalování se stalo v posledních letech aktuální otázkou nejen v České republice, ale v celém světě. Časté jsou analýzy hodnotící difference regionů z hlediska vybraného ekonomického či sociálního ukazatele. Již velmi omezeně však jsou k dispozici studie zabývající se regionálními rozdíly v oblastech, které jsou mnohazměrné a vyžadují použití agregačních metod. Mezi tyto problematiky patří i udržitelný rozvoj a kvalita života. I když se jedná o zcela odlišné koncepce, jejich kvantitativní zachycení je z hlediska využívaných metod často velmi obdobné. V tomto příspěvku je využito dat pro udržitelný rozvoj, nicméně metody zde využitě je možné stejně dobře použít i pro sady ukazatelů hodnotící regiony z hlediska kvality života. Cílem tohoto příspěvku je ukázat možnosti kvantitativního zachycení regionálních diferencí v úlohách vyžadujících agregace většího souboru ukazatelů z různých oblastí, a to pro zemi, která má vzhledem ke své rozloze nízký počet regionů pro úroveň, na níž jsou dostupná dostatečně spolehlivá data.

Data

Pro srovnání regionů jsme zvolili úroveň CZ-NUTS3 územních správních celků, do analýzy tedy vstoupil soubor čtrnácti krajů České republiky. Pro nižší územní celky již nejsou dostupná vhodná data, i tak jsou dostupné časové řady u některých ukazatelů krátké.

Využitá sada ukazatelů byla vytvořena Českým statistickým úřadem a následně publikována v letech 2008 a 2010 (ČSÚ, 2010). Od této doby již nebyla provedena další komparativní studie o udržitelném rozvoji českých regionů. Soubor ukazatelů byl v naší předchozí studii zrevidován s ohledem na dostupnost dat, a to nejen z hlediska dostupnosti ve všech regionech, ale i z hlediska získání maximální délky časové řady. Indikátory byly rozděleny do tří pilířů udržitelného rozvoje - ekonomický (12 ukazatelů), sociální (12) a environmentální (12). Kompletní seznam s detailním vysvětlením změn a úprav oproti původnímu souboru ukazatelů ČSÚ je dostupný v (Fischer a kol., 2013).

Metody

Kompozitní indikátory

Na základě teorie o udržitelném rozvoji byla zvolena dvojstupňová agregace. První krok spočívá v agregaci v rámci pilíře, kdy je pro každý pilíř (ekonomický, environmentální a sociální) provedena samostatná agregace. Výsledná skóre pilířů byla dále agregována do jednoho finálního skóre.

Ukazatele nejsou vyjádřeny ve stejných měrných jednotkách a je tedy třeba nutné nejdříve provést vhodnou transformaci dat, která zároveň pomáhá vyřešit problém s nastavením směru závislosti. Indikátory byly převedeny do směru „vyšší hodnota, lepší pořadí“.

Metodou, která zde byla k transformaci dat využita, je vzdálenost od referenční jednotky. Jako referenční jednotka může být využita nejlepší či nejhorší hodnota v souboru, cílová hodnota či mediánová nebo průměrná hodnota. V našem případě byla využita průměrná hodnota za celou Českou republiku. Pro každý ukazatel je pak hodnota tohoto ukazatele vyjádřena jako poměr k hodnotě referenční jednotky v čase t

$$I_{ij}^t = \frac{x_{ij}^t}{x_{ij}^t}.$$

Metoda zachovává relativní rozdíly mezi jednotkami, z tohoto důvodu je pak snadnější interpretace vzniklých hodnot. Rozptyly jednotlivých ukazatelů zůstávají různé, což na jednu stranu znamená, že není eliminován vliv odlehklých a extrémních pozorování, na druhou stranu lze pak v čase jednoduše sledovat změny ve variabilitě hodnot ukazatele.

Protože není důvod předpokládat, že jeden z pilířů (či ukazatelů) je významnější než další, neaplikujeme žádné váhy, neboli používáme v tomto stupni rovných vah. To samozřejmě zjednodušuje i interpretaci výsledných hodnot. Navíc názor na důsledky korelace mezi jednotlivými ukazateli nemusí být jednotný (Saltelli, 2012). Na jedné straně silná korelace mezi ukazateli může být chápána jako problém, protože může ukazovat na dvojí zahrnutí stejného jevu. Na druhou stranu, silná korelace mezi ukazateli může být znakem měřeného komplexního jevu nebo ukazatele mohou reflektovat nezaměnitelné různé znaky sledovaného jevu a úpravy pak nejsou na místě. Z těchto důvodů byla prozkoumána korelace mezi ukazateli, která nenaznačovala přítomnost nadbytečných ukazatelů v souboru,

Dalším provedeným krokem byla agregace dat. Základní otázkou u agregačních metod je kompenzace mezi ukazateli v agregovaném ukazateli. To znamená možnost, aby nízká hodnota jednoho ukazatele byla kompenzována dostatečně vysokou hodnotou jiného ukazatele. Metody agregace se od sebe liší právě stupněm kompenzace. U lineární agregace pomocí aritmetického průměru je možnost kompenzace konstantní. K nekompenzovatelné agregaci ukazatelů vedou některé nelineární metody odvozené z metod vícekritériálního porovnání, které se však pro účely hodnocení diferencí jeví jako nevhodné z hlediska podávaných výsledků v ordinální škále. Částečně kompenzační metodou agregace je geometrický průměr, který není touto nevýhodou zatížen.

V rámci pilířů bylo k agregaci využito lineární metody agregace. Metoda lineární agregace je nejčastěji užívaná z důvodu snadné srozumitelnosti a jednoduché interpretace. Jelikož lineární agregační metody dovolují plnou kompenzaci, je možné aplikovat tuto metodu uvnitř pilířů, kde lze kompenzaci dovolit. Pro agregaci pilířů je z věcného hlediska celého problému vhodné vycházet z částečně či zcela nekompenzačních metod, tedy geometrického průměru.

Hodnocení regionálních diferencí

Měření regionálních nerovností či jejich zachycení může být teoreticky prováděno pomocí řady různých metod. Některé z těchto metod vedou pouze ke slovním výsledkům zhodnocujících vývoj uvnitř regionů a mezi jednotlivými regiony, jiné jsou založené na statistických metodách. Mezi nejčastěji používané metody v této oblasti patří (Kutscherauer a kol., 2010)

- metoda mezinárodní komparace,
- metody využívající Geografický informační systém,
- poměr maxima a minima,

- metody využívající měř variability – směrodatná odchylka a variační koeficient,
- vícerozměrné statistické metody – metoda hlavních komponent, faktorová analýza a shluková analýza,
- metody reálné konvergence,
- upravený teritoriální Giniho koeficient a
- Theilův index entropie.

V České republice patří mezi často využívané metody především metoda mezinárodní komparace. Výstupem této metody jsou zpravidla regionální analýzy, v jejichž rámci dochází pouze k verbálnímu hodnocení regionálních diferencí. Mezi další využívané metody patří metody využívající Geografický informační systém a ze statistických metod pak především vícerozměrné statistické metody. Nejjednodušší metodou, která je využívána spíše pro orientační výsledky, je poměr maxima a minima. Nevýhodou této metody však je, že může vést k chybným závěrům, neboť jsou využity pouze dvě krajní hodnoty, které mohou být zatíženy odlehlými hodnotami. Je proto vhodné doplnit tuto metodu metodami jinými. Při využití měř variability jsou difference nejčastěji hodnoceny pomocí směrodatné odchylky a variačního koeficientu. Variační koeficient je jako vhodná míra volen především v případě, kdy srovnáváme rozdíly pomocí odlišných charakteristik vyjádřených v různých jednotkách.

V úlohách používajících sociálně ekonomická data poměrně často pracujeme s asymetrickým rozdělením. Pak je důležité, zda je zvolená míra spojena s průměrem a citlivost využití míry na odlehlá pozorování. Tyto vlastnosti splňuje mnohem lépe než výše popsané metody založené na měřách variability Giniho koeficient. Ten je jednak nezávislý na průměru a jednak i názorný ve své interpretaci.

Vzhledem k nízkému počtu pozorování a k výše uvedeným vlastnostem jednotlivých měř nerovností bylo z kvantitativních metod v našem případě hodnocení diferencí v rámci krajů České republiky využito metod využívajících upraveného teritoriálního Giniho koeficientu a Theilova indexu entropie.

Upravený teritoriální Giniho koeficient (Novotný a Nosek, 2006) byl za účelem hodnocení regionálních diferencí vytvořen OECD a využívá vztahu

$$G = \left(\frac{1}{2\bar{x}}\right) \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|,$$

kde x_i , resp. x_j je hodnota sledovaného ukazatele v i tem resp. j tém regionu. Koeficient nabývá hodnoty od 0 do 1, kdy 0 nabude v případě absolutní rovnosti a 1 v případě absolutní nerovnosti.

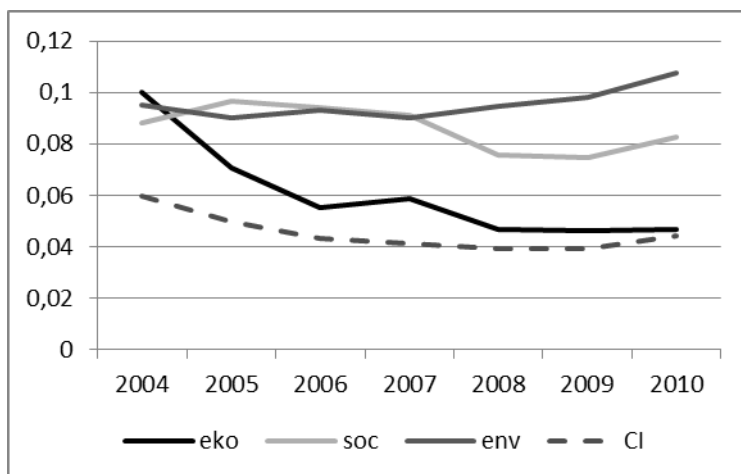
Theilův index entropie nabývá nulové hodnoty v případě absolutní rovnosti a lze ho určit podle vztahu

$$T = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{\bar{x}} \ln \frac{x_i}{\bar{x}},$$

kde x_i je hodnota sledovaného ukazatele v i tém regionu. Výhodou tohoto indexu je možnost jeho rozložení na vnitroskupinovou a meziskupinovou variabilitu. Jinak řečeno lze ho rozložit na složku nerovnosti mezi průměry skupin regionů a složku odpovídající součtu nerovností připadajících na diferenciaci mezi regiony uvnitř jednotlivých skupin regionů, přičemž váhou je zde relativní velikost ukazatele dané skupiny regionů vůči celkovému průměru ukazatele.

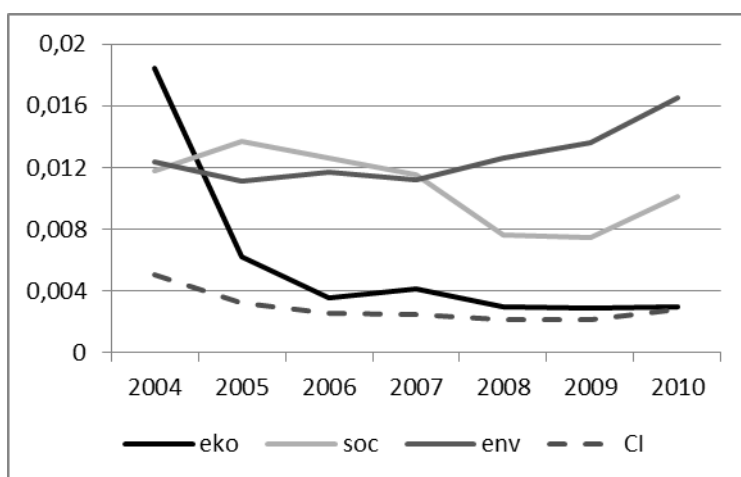
Výsledky

Ukazatele vstupující do výpočtů byly po transformaci agregovány do jednotlivých pilířů (ekonomický, sociální a environmentální) pomocí lineární metody agregace. V dalším kroku došlo k agregaci pilířů pomocí geometrického průměru, jak bylo vysvětleno výše. Pro eliminaci náhodných výkyvů byly použity do výpočtů tříleté průměry hodnot ukazatelů. Pro každý rok byly pro pilíře i kompozit vypočteny upravené teritoriální Giniho koeficienty a Theilovy koeficienty entropie. Výsledky jsou na obrázcích 1 a 2.



Obrázek 1: Upravený teritoriální Giniho koeficient

Zdroj: vlastní výpočty



Obrázek 2: Theilův index entropie

Zdroj: vlastní výpočty

Při pohledu na vývoj měř hodnotících variabilitu u celého kompozitního indikátoru je velmi těžké usuzovat na jednoznačný trend. Z počátku období docházelo k mírnému poklesu, který byl ke konci vystřídán nepatrným nárůstem. Pro vytvoření relevantních závěrů by bylo nutné mít k dispozici delší časovou řadu. Mnohem zajímavější pohled je však na jednotlivé sub-indikátory, tedy na pilíře udržitelného rozvoje.

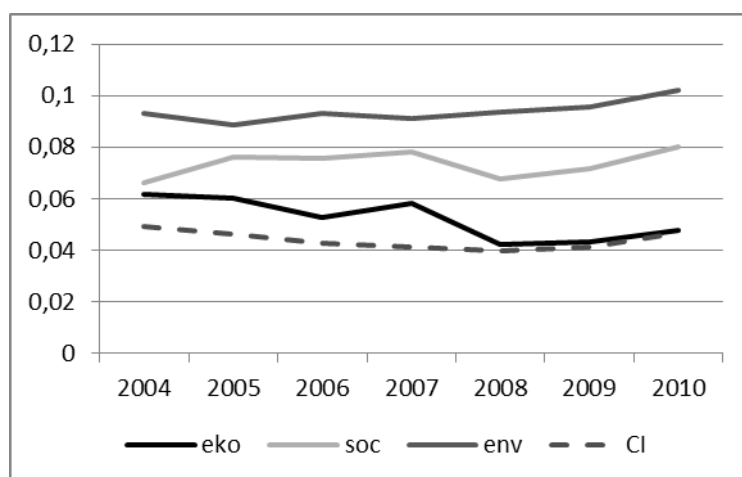
Na základě všech použitých měř lze usuzovat na pokles variability a tedy na snižování regionálních diferencí v rámci ekonomických ukazatelů udržitelného rozvoje. Tento pokles je viditelný především na počátku sledovaného období. Bohužel délka dostupné časové řady neumožňuje zhodnotit předešlý vývoj a říci tak, zdali se v roce 2004 jednalo o krátkodobý výkyv nebo zda byly difference regionů vysoké i před tímto rokem. Opačný trend lze spatřovat u ukazatelů týkajících se environmentální části. V případě skupiny ukazatelů zastupujících sociální oblast nelze spatřovat ve sledovaném období jednoznačný trend.

Vzhledem k tomu, že lze v analýzách udržitelného rozvoje očekávat rozdíly mezi regiony s velkými městy a venkovskými regiony, je vhodné rozdělení regionů alespoň na tyto dvě skupiny. Uvažovány byly dvě možnosti, a to vyjmutí pouze regionu Hl. m. Praha a dále oddělení Hl. m. Prahy a Jihomoravského kraje.

Hl. m. Praha je specifický kraj, který je vymezen hranicemi města a je centrem jak vládních, tak i mnoha vzdělávacích institucí a obchodních společností. Ekonomická síla a vysoce kvalifikované pracovní síly navázané na centrální státní úřady či vědecko-výzkumná a vzdělávací pracoviště jsou hlavním důvodem předního umístění v ekonomickém a sociálním pilíři. Naopak absence ploch zemědělského a přírodního

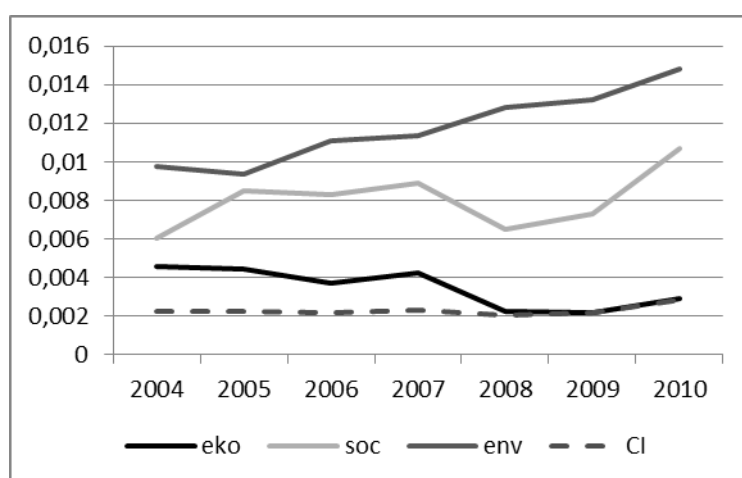
charakteru (toto přirozené zázemí města spadá administrativně již do Středočeského kraje) způsobuje umístění tohoto regionu na poslední příčce v pilíři environmentálním. V určité míře lze spatřovat podobnost v umístění v prvních dvou pilířích i u Jihomoravského kraje, jehož správním centrem je město Brno, byť v porovnání s Prahou je Brno centrem spíše regionálního významu. Podobnost pak rozhodně není u umístění ve třetím pilíři, neboť území tohoto kraje nezahrnuje pouze samotné město (jako je tomu v případě Prahy), ale též přírodně velmi rozmanité regiony jižní Moravy.

Z výše uvedených dvou variant byla vybrána varianta druhá, kdy po odebrání obou dvou krajů došlo k výraznému poklesu regionální diferenciace naznačující zatížení výsledků přítomností obou krajů v souboru. Výsledné hodnoty měr diferencí pro zbylých 12 regionů jsou na obrázcích 3 a 4.



Obrázek 3: Upravený teritoriální Giniho koeficient

Zdroj: vlastní výpočty

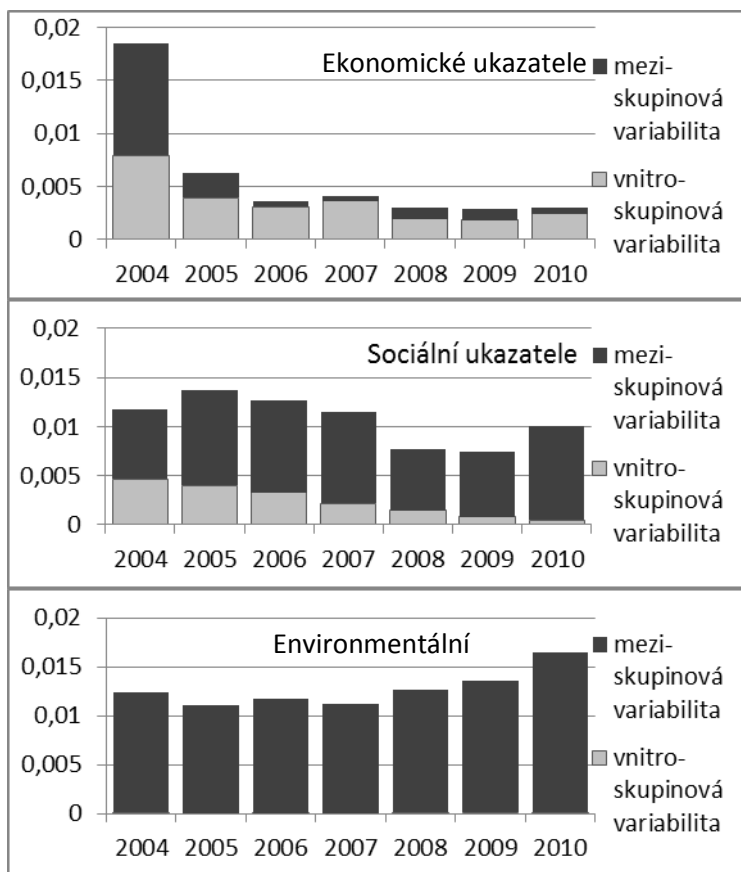


Obrázek 4: Theilův index entropie

Zdroj: vlastní výpočty

Po odebrání regionů z analýzy je vidět vyrovnaní křivek a především zde již není tak znatelný pokles u ekonomických ukazatelů. I tak lze ale říci, že v daném časovém období docházelo k poklesu regionálních diferencí v oblasti ekonomického pilíře udržitelného rozvoje a naopak k jejich mírnému růstu v pilíři environmentálním. U sociálního pilíře došlo k poklesu celé křivky. Ani po odstranění statisticky odlehklých regionů však nelze usuzovat na trend ve sledovaném období. Velmi dobře je ale viditelné, že k největším regionálním diferencím dochází v České republice z hlediska udržitelného rozvoje v environmentálních ukazatelích a naopak k nejmenším v oblasti ekonomické.

Jak již bylo řečeno výše, výhodou Theilova indexu entropie je jeho rozložitelnost na variabilitu uvnitř skupin regionů a mezi těmito skupinami. Toho lze využít pro dvě vytvořené skupiny regionů.



Obrázek 5: Rozklad Theilova indexu entropie na meziskupinovou a vnitroskupinovou variabilitu

Zdroj: vlastní výpočty

Na grafu znázorňujícím rozklad variability u ekonomických ukazatelů je velmi dobře viditelná vysoká meziskupinová variabilita, která se značnou mírou podílí na výši obou indexů na obrázku 1 resp. 2. Za povšimnutí stojí, že v následujících letech je většina variability mezi regiony tvořena vnitroskupinovou složkou a nelze tedy usuzovat na velké rozdíly mezi regiony s velkými městy a zbylé regiony z hlediska ekonomického pilíře udržitelného rozvoje. U sociálního pilíře je v průběhu času zřetelný pokles vnitroskupinové variability, který je však dán i sbližováním obou regionů s velkými sídly. Meziskupinová variabilita je však po celou dobu velmi vysoká a lze tedy usuzovat na značné rozdíly mezi oběma skupinami regionů. Největší podíl meziskupinové variability lze spatřovat u pilíře environmentálního. To je dáno zcela přírodně odlišným charakterem městských regionů. Jihomoravský kraj je však v tomto ohledu regionem, který některými svými částmi odpovídá spíše skupině krajů bez velkých sídelních center.

Závěr

V příspěvku bylo využito kompozitních indikátorů k hodnocení regionálních diferencí pro mnohazměrné úlohy vyžadující zhodnocení větší skupiny ukazatelů z různých hledisek. Vzhledem k dostupnosti dat byla analýza provedena na krajské úrovni (NUTS 3), která obsahuje 14 krajů. Nedostupnost dat pro nižší územně správní celky komplikuje či téměř znemožňuje využití některých statistických metod z důvodu velmi nízkého počtu pozorování, pro naše účely tedy bylo využito upraveného teritoriálního Giniho koeficientu a Theilova indexu entropie. Druhá z těchto měr navíc umožňuje využití variability mezi jednotlivými regiony na meziskupinovou a vnitroskupinovou variabilitu, čehož bylo využito při hodnocení diferencí po rozdělení krajů na dvě skupiny, skupinu regionů s velkými sídly a zbylé regiony.

Na základě výsledků prezentovaných v příspěvku lze říci, že díky použitým mírám byl v dostupném období vysledován trend sbližování či oddalování krajů České republiky v rámci pilířů udržitelného rozvoje. Tento trend je viditelný především u pilířů ekonomického (sbližování krajů) a environmentálního (oddalování krajů). Zajímavé jsou výsledky rozkladu variability, které ukazují na velké rozdíly mezi regiony s velkými sídly a zbylými regiony v rámci pilíře sociálního i environmentálního.

Zde ukázané metody jsou možností k přístupu hodnocení regionálních diferencí v úlohách výše popsaného typu, kdy nelze použít většinu statistických metod. Bohužel absence delších časových řad znemožňuje tvorbu relevantnějších závěrů, i tak bylo ale možné za dané časové období vysledovat vývoj rozdílů v regionech České republiky z hlediska udržitelného rozvoje a jeho pilířů a ukázat možné využití použitých měr pro hodnocení skupin regionů.

Literatura

ČESKÝ STATISTICKÝ ÚŘAD, 2010. *Vybrané oblasti udržitelného rozvoje v krajích České republiky*. Praha: Český statistický úřad

FISCHER J., PETKOVÁ L., HELMAN K., KRAMULOVÁ J., ZEMAN J., 2013. *Sustainable development indicators at the regional level in the Czech Republic*. Statistika, r. 50 č. 1

KUTSCHERAUER A. a kol., 2010. *Regionální disparity - Disparity v regionálním rozvoji země - pojetí, teorie, identifikace a hodnocení*. Ostrava: VŠB-Technická univerzita Ostrava

NOVOTNÝ J., NOSEK V., 2006. *Regionální dimenze sociálně-ekonomických nerovností v Česku: pojetí, měření, empirie*. Sborník příspěvků z XXI. sjezdu České geografické společnosti

SALTELLI A. 2012. *Composite Indicators: An introduction*. Paper presented at the 10th JRC Annual Seminar on Composite Indicators.

JEL Classification: Q01, R11, R59.

Dedikace: Příspěvek je zpracován jako součást výzkumného projektu Vysoké školy ekonomické v Praze č. 11/2012 Konstrukce a verifikace indikátorů udržitelného rozvoje ČR a jejích regionů.

Summary

Statistical analysis of regional differences of sustainable development and quality of life

Czech Republic passes in the last 30 years of huge economic and social changes and attention in this context is also given on the development of individual regions. The question is, if this transition will lead to a convergence of individual region or vice versa.

The paper is focused on the possibilities of assessing regional differences in areas that are often described quantitatively hardly describable phenomena leading to a multi-dimensional tasks. Indicator's sets present plenty of data sets contained in these respects confusing and restrictive. With using this indicator's sets disappear the simplicity and accuracy of the result. The trend in recent years is find another way, for example the composite indicators. A major advantage of composite indicators is the possibility of summarization, which allowing easy comparison of regions.

Part is focused on methods for monitoring of regional differences aggregate indicators involving large amounts of disparate indicators in the Czech Republic. The methods are illustrated on data for sustainable development, however, is very similar to the use in the evaluation of quality of life. The quality of life is despite the differences in their conception very often similar to sustainable development in the quantitative approach.

Key words: Regional differences, Composite indicators, Regional analysis, Sustainable development.

Pracovní migrace v České republice

Martina Šimková

martina.simkova@vse.cz

Doktorand oboru statistika

Školitel: doc. Ing. Jitka Langhamrová, CSc. (langhamj@vse.cz)

Abstrakt: Česká republika, stejně jako většina ekonomicky vyspělých evropských zemí, se dnes potýká s problémem stárnutí populace. Současný demografický vývoj a z něho odvozené populační prognózy ukazují na neustálý budoucí růst podílu lidí v důchodovém věku. Přibývání počtu starších osob je pochopitelně doprovázeno klesajícím podílem produktivní populace. Jednou možností, jak zmírnit důsledky stárnutí, je aktivní migrační politika, neboli posílení produktivní populace pracovní silou ze zahraničí. Tento článek se zabývá otázkou, zda migrace má na stárnutí populace České republiky pozitivní či negativní vliv. Hlavním cílem je analyzovat vývoj a současný stav imigrace a emigrace za prací, zhodnotit chování migrantů na cizím trhu práce a rozhodnout, zda prostřednictvím transferů výdělků cizinců (tzv. remitencí) odplouvá (resp. připlouvá) významné množství prostředků z ČR (resp. do ČR). Pro ilustraci významu posílání remitencí na ekonomiku České republiky je přepočítána výše odeslaných remitencí k hrubému domácímu produktu, tak jak je to obvykle prezentováno na mezinárodním poli. Na závěr je také provedeno mezinárodní srovnání se zeměmi, pro které migrace z pohledu odeslaných remitencí představuje významnější otázku.

Klíčová slova: stárnutí populace, trh práce, pracovní migrace, remittance.

Úvod

Stárnutí populace je charakteristickým rysem současného a budoucího demografického vývoje České republiky i dalších rozvinutých evropských zemí. Vývoj úmrtnosti v České republice je v poslední době na celkem stabilní úrovni, střední délka života mírně roste. Naopak v případě plodnosti a migrace nastaly v posledních zhruba dvaceti letech poměrně výrazné změny v trendech vývoje. Ve druhé polovině devadesátých let minulého století došlo k prudkému poklesu plodnosti, která poté sice mírně rostla, nicméně současný vývoj naznačuje stagnaci plodnosti na úrovni, která nezaručuje prostou reprodukci obyvatelstva. Úroveň plodnosti se sice může mírně zvýšit v příštích desetiletích, ale není pravděpodobné, že by v České republice i v ostatních zemích dosáhla úrovně prosté reprodukce. Lze tedy předpokládat, že význam imigrace pro udržení početního stavu populace nadále poroste. Imigrace hraje roli především v zachování stabilní úrovně pracovní síly, úbytek osob v produktivním věku je částečně kompenzován přílivem ekonomicky aktivních imigrantů. Lidé v poproduktivním věku budou tedy závislí i na ekonomické aktivitě migrantů. V souvislosti se stárnutím populace bude v tomto článku diskutována otázka, zda aktivní migrační politika může zmírnit jeho důsledky a jaké demograficko-ekonomické dopady může mít migrace na Českou republiku.

Je ovšem třeba vzít v úvahu, že jedním z důvodů migrace za prací je finanční zabezpečení členů rodiny v zemi původu, mnoho migrantů totiž posílá část vydělaných peněz svým rodinám ve formě remitencí. To má sice jistě pozitivní vliv na ekonomickou situaci chudších zemí, ze kterých migranti přicházejí, nicméně ekonomice České republiky tak odplouvají prostředky do zahraničí. Migrace ale samozřejmě existuje i v opačném směru, i občané České republiky často migrují do zahraničí za prací a posílají část výdělků zpět do země. Je tedy možné, že odliv prostředků odesláním do zahraničí je částečně kompenzován přílivem remitencí ze zahraničí.

Cílem tohoto článku je představit chování migrantů na českém trhu práce, analyzovat vývoj imigrace a emigrace za prací do (z) České republiky od roku 1993, a rozhodnout, zda remittance významně ovlivňují ekonomiku České republiky. Pro tyto analýzy využívám data Českého statistického úřadu (ČSÚ) o počtech migrantů ve struktuře podle délky pobytu, ekonomické aktivity a dle země původu (resp. cílové země).

Názory na význam migrace

Pracovní migrace patří mezi často diskutovaná témata v souvislosti se stárnutím populace, protože má pozitivní i negativní dopady na hostitelskou ekonomiku. Většina studií naznačuje, že imigrace přispívá k hospodářskému růstu v hostitelské zemi tím, že zmírní důsledky stárnutí populace, které způsobuje nerovnováhu mezi velikostí produktivní a poproduktivní populace. Zejména s úlohou systému sociálního zabezpečení budou fiskální důsledky imigrace pozitivní, protože imigranti zde platí sociální pojištění, jehož sazby by se jinak musely výrazně zvyšovat kvůli rostoucímu počtu osob v důchodovém věku a klesajícímu počtu osob v produktivním věku.

Počty migrantů potřebné k vyrovnání poklesu produktivní populace jsou ale vysoké, a jestli je větší počet imigrantů v rámci možností státu, závisí do značné míry na sociálních, ekonomických a politických okolnostech dané zemi nebo regionu (United Nation, 2000). Dle studie J. Equetera (2002) by bylo potřeba do Evropských zemí přilákat do roku 2050 téměř 80 milionů pracovníků k udržení velikosti potenciální pracovní síly a asi 700 milionů migrujících pracovníků k udržení vyváženého poměru produktivní a neproduktivní populace. Dle J. Muyskena (2008) imigrace sama o sobě nemůže udržet stabilitu produktivní populace, je také potřeba zvýšit míru účasti na trhu práce, zejména ve vyšším věku. Velikost produktivní populace lze také udržet postupným prodlužováním doby ekonomické aktivity prostřednictvím zvyšování hranice pro odchod do důchodu. Dle studie United Nation (2000) ve většině případů může být zachování poměru produktivní a poproduktivní populace na současné úrovni docíleno zvýšením hranice důchodového věku do zhruba 75 let. V České republice v rámci důchodové reformy probíhá zvyšování důchodového věku. Současný cílový stav po roce 2030 stanovuje důchodový věk na 65 let u mužů a u žen na 62 až 65 letech podle počtu vychovaných dětí. Ke sjednocení důchodového věku tak dojde v roce 2041 pro pojištěnce narozené v roce 1975 (MPSV, 2011a). Některé studie (např. Park, 2007) dokonce upozorňují na skutečnost, že imigranti mohou zabírat pracovní příležitosti rodilým pracovníkům, a to zejména těm, kteří jsou mladí a mají nízké dovednosti. U širší veřejnosti také bývají často rozšířené obavy z příliš vysoké porodnosti přistěhovalců.

Remittance jako fenomén migrace

Česká republika patří již nějakou dobu mezi imigrační země, kam přichází každoročně množství migrantů za prací. Chování migrantů závisí jednak na délce a účelu pobytu v České republice a jednak na vazbě na mateřskou zemi. Někteří cizinci zde totiž pracují především proto, aby mohli posílat zde vydělané peníze domů svým rodinám ve formě tzv. remitencí. Obecně by se tedy dalo říci, že čím bude imigrace vyšší, tím bude ekonomice České republiky odplouvat více prostředků do zahraničí. Pro země, kde vyplacené remittance tvoří velkou část hrubého domácího produktu, může imigrace způsobovat problém.

Remittance jsou definovány jako převody peněz migrantů do země původu vycházející z jejich dočasných nebo trvalých příjmů (IMF, 2009). Pro kvantifikaci celkového množství remitencí je nutné znát počet migrantů a výši jejich příjmů. Odhad příjmů a remitencí je založen na chování cizinců. Primárním krokem ve výpočtech je zjištění počtu a struktury cizinců na území České republiky a občanů České republiky v zahraničí. Od počtu a struktury cizinců se odvíjí finanční chování (příjmy, výdaje, úspory, remittance).

K problematice cizinců je možno přistupovat různě, ale pro vyjádření jejich vlivu na ekonomiku je nezbytné se opřít o standard národního účetnictví ESA 1995. V národních účtech jsou cizinci rozlišováni primárně podle délky pobytu na:

- 1) Rezidenty – ekonomický subjekt, sídlící a působící na území ČR alespoň jeden rok nebo déle,
- 2) Nerezidenty – ekonomický subjekt, sídlící a působící na území ČR méně než jeden rok (příhraniční pracovníci – tzv. pendleři, sezónní pracovníci, studenti studující v ČR, cizinci pracující na českých zastupitelských úřadech v zahraničí).

Rozdělení cizinců na rezidenty a nerezidenty je klíčové pro správné zachycení a vyčíslení souvisejících finančních toků. Z pohledu Čechů v zahraničí jsou rezidenti pro Českou republiku ekonomické subjekty, které působí na území cizího státu méně než jeden rok (příhraniční pracovníci, sezónní pracovníci, studenti studující v zahraničí a Češi pracující na cizích zastupitelských úřadech v České republice), naopak ekonomické subjekty sídlící v zahraničí déle než jeden rok, jsou pro Českou republiku nerezidenty.

Dále je třeba vzít v úvahu členění dle ekonomické aktivity:

- 1) Ekonomicky aktivní – zaměstnanci (legální, nelegální) x podnikatelé,
- 2) Ekonomicky neaktivní – studenti x ostatní.

Podnikatelé jsou podle současných pravidel národního účetnictví považováni pouze za rezidenty. Za třetí se migranti sledují dle země původu (resp. dle cílové země). Toto rozdělení je důležité, protože lidé z chudších zemí mají mírně odlišné chování.

Hlavní datové zdroje

Počet a struktura cizinců bohužel není sledována jednotně, neexistuje žádná databáze, která by shromažďovala veškeré informace o cizincích. Východiskem pro zjištění počtu cizinců na území České republiky jsou tak tři základní nezávislé administrativní zdroje:

- 1) Ministerstvo vnitra (MV) – Cizinecká policie eviduje celkový počet cizinců dle délky pobytu a země původu,
- 2) Ministerstvo práce a sociálních věcí (MPSV) – Úřady práce evidují počty vydaných pracovních povolení,
- 3) Ministerstvo průmyslu a obchodu (MPO) – evidence živnostenských oprávnění.

Odhad počtu cizinců zaměstnanců a podnikatelů se provádí kombinací všech dostupných datových zdrojů. Znalost celkového počtu cizinců v členění podle délky pobytu od MV umožňuje rozdělení cizinců na rezidenty a nerezidenty. Počet ekonomicky neaktivních obyvatel vychází z rozdílu celkového počtu cizinců rezidentů, počtu legálně zaměstnaných obyvatel (MPSV) a podnikatelů (MPO).

Podobný postup je aplikován na výpočet ekonomicky neaktivních nerezidentů. Nejprve je odhadnuta nelegální zaměstnanost cizinců. Odhady nelegálně zaměstnaných cizinců jsou založeny na výsledcích inspekci provedených kontrol u zaměstnavatelů zaměstnávajících cizince, ve spolupráci s celními úřady, orgány Služby cizinecké policie a s oblastními inspektoráty práce (MPSV, 2011b). Nelegálním zaměstnancem se nejčastěji stane občan třetí země, který čeká na vydání rozhodnutí příslušným státním orgánem po dobu delší než 3–6 měsíců. Podle zákona o zaměstnanosti má žadatel o mezinárodní ochranu povolen vstoup na trh práce až po uplynutí 1 roku ode dne podání žádosti.

Mezi ekonomicky neaktivní nerezidenty se zahrnují zahraniční studenti a ostatní neaktivní osoby (např. ženy v domácnosti). Počty zahraničních studentů studujících v České republice jsou získány z Ústavu pro informace ve vzdělávání (ÚIV)¹. Počet ostatních ekonomicky neaktivních nerezidentů se vypočítá jako rozdíl celkového počtu nerezidentů a ekonomicky aktivních nerezidentů (legálně + nelegálně zaměstnaných) a zahraničních studentů.

Současné zdroje dat o počtech Čechů sídlících a působících v zahraničí jsou dva:

- 1) Výběrové šetření pracovních sil (VŠPS) – počet občanů ČR pracujících v zahraničí po dobu kratší než 1 rok nebo dojíždějících do zahraničí za prací,
- 2) Všeobecná zdravotní pojišťovna (VZP) – počet občanů ČR pobývajících v zahraničí.

VŠPS poskytuje údaje o počtu českých občanů rezidentů pracujících v zahraničí. Pro bližší specifikaci dlouhodobých pobytů využívá ČSÚ podklady VZP (jako nejvýznamnější zdravotní pojišťovny v ČR), konkrétně údaje o počtech osob s trvalým pobytem v České republice, které se v souvislosti s migrací odhlásily (přihlásily) ze (do) systému veřejného zdravotního pojištění. Tato data jsou dostupná v členění podle věku migrantů a délky pobytu v cizině, což umožňuje odhad celkového počtu Čechů rezidentů a nerezidentů.

Metodika odhadu remitencí

Objem zaslaných prostředků cizinci (tzv. remitencí) se liší dle ekonomické aktivity a délky pobytu cizince. Remitence jsou velmi špatně měřitelné, existuje velice málo přímých informací, odhad výše remitencí je proto dosud prováděn pomocí kvalifikovaných odhadů. Český statistický úřad odhaduje výši

¹ Ústav pro informace ve vzdělávání byl z rozhodnutí ministra školství, mládeže a tělovýchovy k 31.12.2011 ZRUŠEN. Od 1.1.2012 přebírá agendy ÚIV spojené se sběrem a zpracováním dat MŠMT. <http://www.msmt.cz/vzdelavani/socialni-programy/dokumenty-6>

remitencí u cizinců rezidentů přímo s využitím výzkumného projektu Sociologického ústavu Akademie Věd ČR (dále SOÚ), který byl proveden v roce 2010². Projekt byl zaměřen na migraci pracovních sil, příjmy, výdaje na konečnou spotřebu, úspory a remittance několika národností v České republice (Ukrajina, Vietnam, Rusko, Moldávie a země bývalé Jugoslávie). Výsledky tohoto projektu byly použity pro odhad peněžních prostředků a naturálních remitencí cizinců z pěti výše uvedených zemí. Chování imigrantů z jiných zemí je expertně odhadnuto na základě podobnosti s vybranými zeměmi. Pomocí dat z tohoto projektu byl odhadnut podíl remitencí a podíl věcných darů³ z příjmů, které cizinci posílají do země původu svým rodinám a příbuzným. Výše remitencí zaměstnaných cizinců rezidentů a podnikatelů se pak rovná:

$$R_i^{(zam)} = NW_i \cdot (r + g) \quad (1)$$

$$R_i^{(pod)} = NLB_i \cdot (r + g), \quad (2)$$

kde

$R_i^{(zam)}$ remittance odeslané rezidentskými zaměstnanci do země i ,

$R_i^{(pod)}$ remittance odeslané podnikateli do země i ,

NW_i čistá mzda

NLB_i schopnost(+)/potřeba(-) financování podnikatelů,

r procentní podíl peněžních remitencí z příjmu,

g procentní podíl darů z příjmu a

i index dané země.

Dle standardu národního účetnictví nerezidenti nevytvářejí úspory v České republice, nebo jejich úspory jsou pouze dočasné a budou dříve či později použity v jejich domovské zemi. Při odhadu remitencí nerezidentů tedy předpokládám rovnost mezi remittencemi a úsporami. To znamená, že všechny získané prostředky zbývající po odečtení výdajů na konečnou spotřebu, jsou posílány zpět do domovské země. Pro zjištění úspory cizinců jsou od výše čistých mezd odečteny výdaje na konečnou spotřebu. Tyto výdaje jsou odhadnuty pomocí struktury výdajů českých domácností na konečnou spotřebu dle klasifikace COICOP⁴, které se pak následně individuálně opravují pro každou skupinu (Ondruš, 2009). Remittance se tedy rovnají čisté mzdě snížené pouze o jejich spotřebu na území České republiky:

$$R_i^{(nerez)} = S_i = NW_i - CE_i, \quad (3)$$

kde

$R_i^{(nerez)}$ remittance odeslané nerezidentskými zaměstnanci do země i ,

S_i úspory,

NW_i čistá mzda,

CE_i výdaje na konečnou spotřebu a

i index dané země.

Základem pro výpočet úspor a remitencí je čistá mzda. Čistá mzda musí být členěna podle typu činnosti, záležitosti, jestli je cizinec zaměstnán legálně nebo nelegálně, nebo podniká. ČSÚ odhaduje hrubé mzdy na základě databáze spravované soukromou společností Trexima (ISPV, 2011). Zde jsou k dispozici údaje o průměrných mzdách zaměstnanců vybraných občanství pracujících v České republice. Bohužel tento průzkum nerozlišuje délku a účel pobytu a všechna občanství. Vybrané získané průměrné mzdy jsou

² Více viz Leontiyeva Y., Tollarová B. 2011

³ Cenné předměty, jako elektronické zboží, drahé léky, oblečení, šperky, apod.

⁴ Klasifikace CZ-COICOP slouží k zařazení všech druhů individuální spotřeby (zboží, služeb apod.) podle účelu. Je používána k identifikaci výdajů na individuální spotřebu třemi institucionálními sektory: domácnostmi, neziskovými institucemi sloužícími domácnostem a vládními institucemi.

použity pro odhad průměrných mezd ve všech ostatních zemích. Čistá mzda je rovna mzdám a platům sníženým o příspěvky na sociální a zdravotní pojištění a daně z příjmů:

$$NW_i = WS_i - (TI_i + ESC_i), \quad (4)$$

kde

NW_i čistá mzda zaměstnance,

WS_imzdy a platy,

TI_i daň z příjmu,

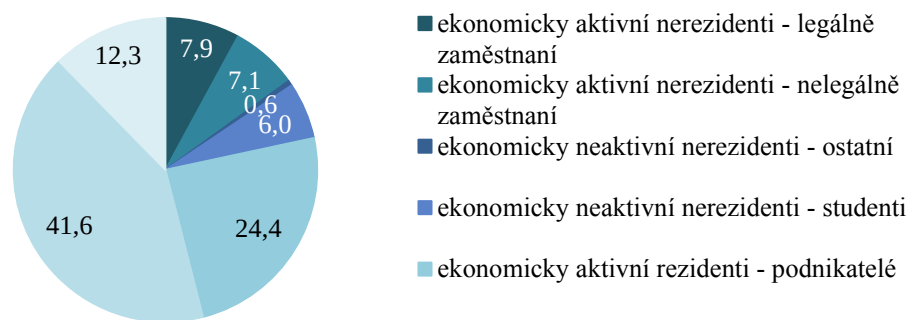
ESC_isociální příspěvky zaměstnance a

iindex dané země.

Výše uvedený vzorec (4) se používá pro výpočet čisté mzdy pouze pro legálně zaměstnaného nerezidenta a zaměstnaného rezidenta, protože nelegálně zaměstnaný cizinec neodvádí žádné příspěvky ani daně. Jeho čistá mzda je rovna hrubé mzdě. Ovšem lze předpokládat, že nelegální zaměstnanec dostane od zaměstnavatele nižší mzdu přibližně o výši daní a příspěvků (tj. v průměru o 25 %). Čistá mzda cizinců podnikatelů se odhaduje pomocí schopnosti/potřeby financování českých podnikatelů a tvoří čistý příjem, který mají podnikatelé k dispozici pro jejich osobní výdaje a pro případný převod peněz do země původu (Ondruš, 2009). Schopnost/potřeba financování cizinců podnikatelů je expertně odhadnuta pro jednotlivé druhy zemí, zpravidla platí, že schopnost/potřeba financování cizinců podnikatelů jsou mírně vyšší než v případě českých podnikatelů.

Struktura cizinců v České republice

V České republice je zatím počet imigrantů relativně nízký. Podle odhadů s využitím různých administrativních zdrojů bylo v České republice v roce 2012 přibližně 562 tisíc cizinců, což v relativním vyjádření tvoří 5 % celkového počtu obyvatel. Z toho přibližně 80 % tvořili pracující cizinci (zaměstnanci rezidenti – 42 %, podnikatelé – 24 % a zaměstnanci nerezidenti – 15 %). Hlavním důvodem pro migraci do České republiky je tedy bezpochyby pracovní migrace (viz obrázek 1). V poslední době se ovšem zvyšuje počet cizinců, jejichž účelem pobytu je sjednocení rodiny či usídlení, což dokazuje skutečnost, že v roce 2012 bylo v České republice 13 % ekonomicky neaktivních cizinců. Podíl cizinců pobývajících za účelem studia se také zvyšuje, dnes tvoří cizí studenti 6 % z celkového počtu migrantů, v roce 2012 to bylo téměř 34 tisíc studentů.



Obrázek 35 Složení cizinců v ČR dle délky pobytu a ekonomické aktivity (2012)

Zdroj: ČSÚ, vlastní výpočty

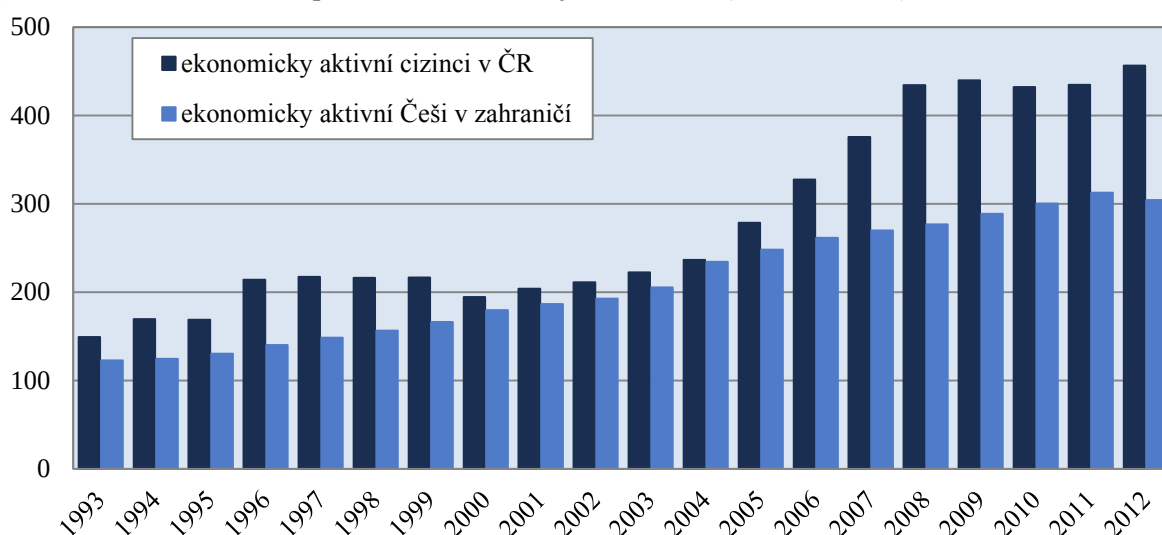
Mezi cizinci bylo v roce 2012 56,8 % mužů. Hlavní důvod migrace do České republiky také potvrzuje věkové složení cizinců, v ekonomicky aktivním věku (15–64 let) se nacházelo 87 % cizinců (ČSÚ, 2013). Zajímavé je národnostní složení pracujících cizinců. Následující tabulka zobrazuje pět nejpočetnějších národností v České republice z hlediska délky pobytu. Nejvíce osob přichází z Ukrajiny a ze Slovenska. Pro skupinu migrantů z Ukrajiny je charakteristický vysoký počet nelegálně zaměstnaných pracovníků. Větší procento osob z Ukrajiny zde také podniká, stejně jako imigranti z Vietnamu. Cizinci v postavení zaměstnanců pracují především ve zpracovatelském průmyslu, stavebnictví, velkoobchodu a maloobchodu.

Tabulka 5 Struktura pracujících cizinců v roce 2012

Zdroj: ČSÚ

2012	Zaměstnanci celkem	Rezidenti			Nerezidenti		
		Celkem	Zaměstnanci	Podnikatelé	Celkem	Legálně zaměstnaní	Nelegálně zaměstnaní
Celkem	456 310	371 613	234 379	137 234	84 697	44 606	40 091
Ukrajina	126 688	84 558	34 154	50 404	42 130	19 082	23 048
Slovensko	120 035	114 140	102 214	11 926	5 895	5 895	0
Vietnam	42 255	39 835	2 703	37 132	2 420	1 096	1 324
Rusko	26 528	19 413	12 286	7 127	7 115	3 227	3 888
Polsko	21 821	20 755	18 972	1 783	1 066	1 066	0

Růst počtu cizinců v České republice je patrný již od roku 1993. Za uplynulých dvacet let vzrostl celkový počet cizinců ze 159 tisíc na 563 tisíc (nárůst o více než 400 tisíc). Po celé zkoumané časové období tvoří ekonomicky aktivní cizinci okolo 80 % z celkového počtu cizinců. K dočasnému poklesu počtu povolení k pobytu cizinců došlo v roce 2000, a to důsledkem změny zákona č. 326/1999 Sb., o pobytu cizinců na území České republiky (MZV, 2013). Po vstupu České republiky do EU se změnila struktura zaměstnanosti cizinců, občané EU již nepotřebují pracovní povolení a jejich zaměstnavatelé mají pouze povinnost informovat úřad práce, že zaměstnávají občana EU (MPSV, 2011b).



Obrázek 36 Vývoj počtu pracujících cizinců v ČR a Čechů v zahraničí (tisíce osob)

Zdroj: ČSÚ, vlastní výpočty

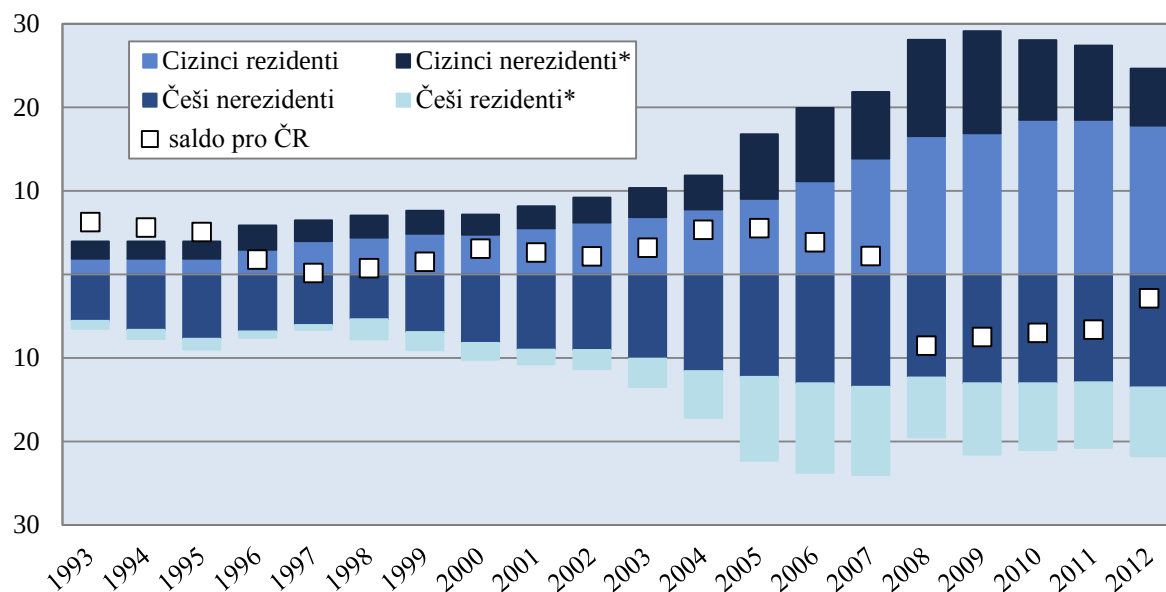
Zájem cizinců o práci v České republice dnes neklesá ani přes nepříznivý vývoj na trhu práce v posledních letech. Přestože Česká republika doposud nepatří, ve srovnání s ostatními členskými zeměmi EU, mezi významné imigrační země, cizinci jsou považováni za potenciální přínos pro ekonomiku České republiky.

Na druhou stranu ale také mnoho občanů České republiky pracuje v zahraničí, v období let 2000 – 2005 se dokonce dle odhadů ČSÚ velmi blížil počet cizinců v ČR a Čechů v zahraničí. Se vstupem do EU v roce 2004 se počet cizinců v ČR a počet Čechů v zahraničí zvýšil.

Vývoj objemu remitencí v ČR

Pracovní migrace hraje důležitou roli v udržování stabilní úrovně pracovní síly. Jedním z projevů migrace jsou ovšem remittance. Transfery peněz migrantů do zemí původu mají jistě pozitivní vliv na situaci chudších ekonomik, ze kterých migranti přicházejí, nicméně české ekonomice tak odplouvají prostředky do zahraničí.

Občané České republiky pracují v cizině spíše krátkodobě (viz obrázek 3). Pokud zůstanou déle než rok, výše částek zasílaných domů klesá, protože si začínají budovat nový domov tam. U cizinců v České republice je to však naopak, čím déle zůstávají, tím více své příbuzné podporují. Dá se z toho usuzovat, že vazba cizinců na domov je podstatně vyšší než u Čechů.



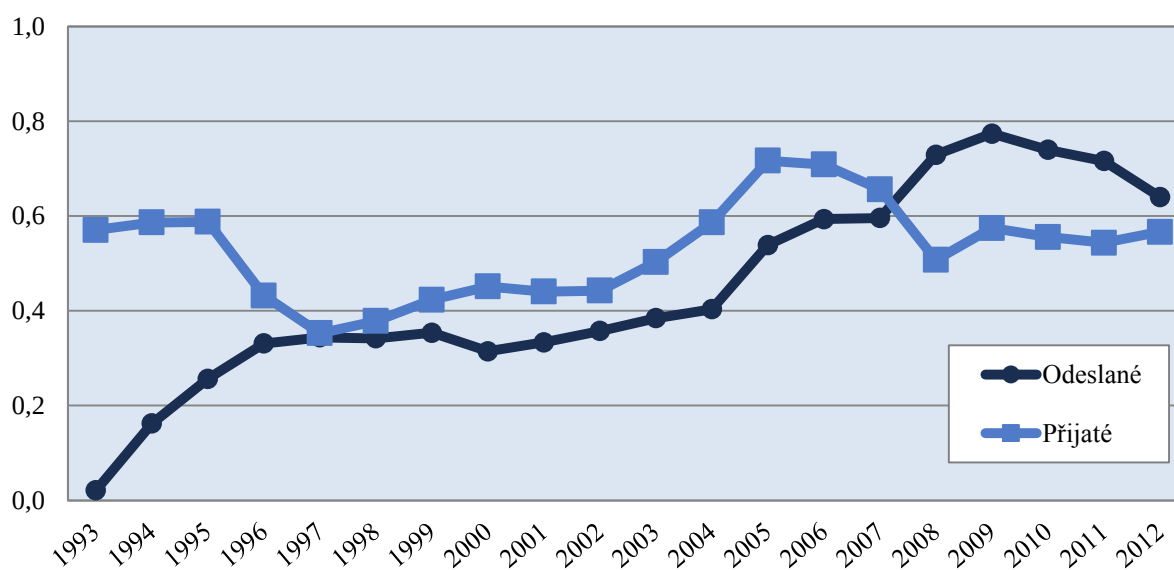
Obrázek 37 Objem odeslaných a přijatých remitencí v ČR (mld. Kč)

Zdroj: The World Bank 2013, ČSÚ, vlastní výpočty

* výše remitencí je rovna celkové úspoře

Celkový objem prostředků odeslaných z ČR dosáhl v roce 2012 24,6 mld. Kč, přijaté remittance byly rovny téměř 22 mld. Kč. Do roku 2007 přijaté remittance převažovaly nad odeslanými. Od roku 2008 převládají odeslané remittance, rozdíl však není vysoký. Ve srovnání s evropskými zeměmi je dnes Česká republika na 31. místě v žebříčku objemu odeslaných remitencí.

Pro lepší představu o vlivu posílání remitencí na ekonomiku ČR jsem přepočítala výši odeslaných a přijatých remitencí k výši HDP. V roce 2012 tvořily odeslané remittance přibližně 0,64 % HDP a přijaté remittance 0,57 % HDP. Počátkem 90. let měly odeslané remittance bezvýznamný (téměř nulový) podíl na HDP.

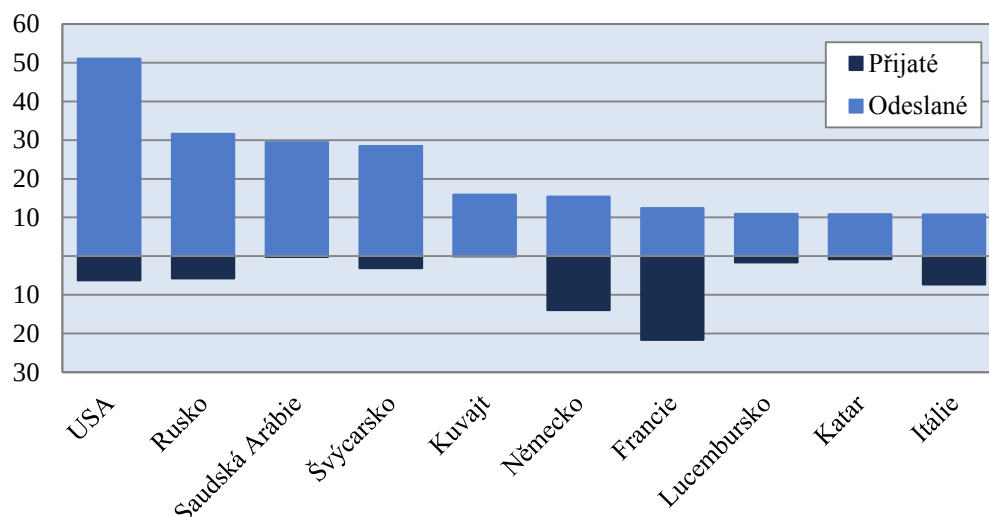


Obrázek 38 Podíl odeslaných a přijatých remitencí na HDP (%)

Zdroj: ČSÚ, vlastní výpočty

Mezinárodní srovnání remitencí

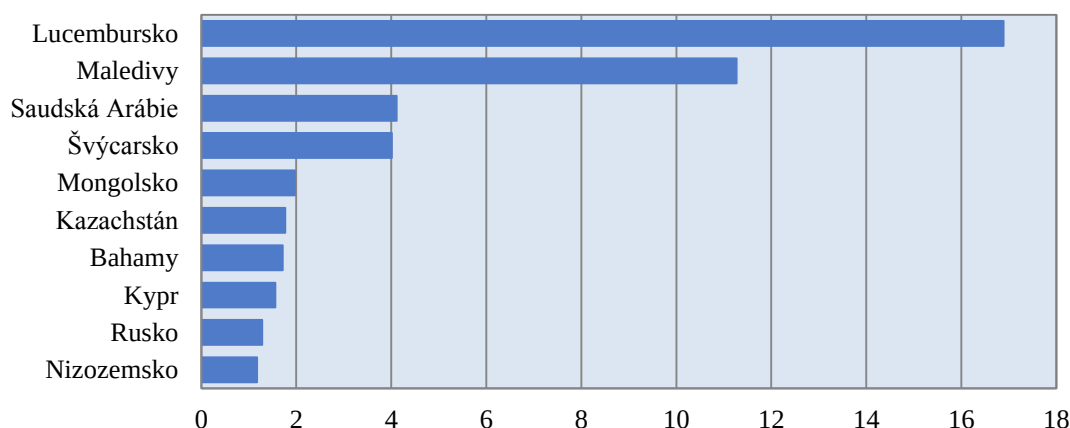
Pro představu o vlivu migrace a remitencí na českou ekonomiku uvádím srovnání s několika zeměmi světa, pro které migrace představuje problém z hlediska odeslaných remitencí (viz obrázek 5). Největší množství remitencí v roce 2012 bylo odesláno ze Spojených států amerických (51 mld. USD) a z Ruska (32 mld. USD). Do těchto dvou zemí se v opačném směru se vrátila pouze desetina prostředků ve formě přijatých remitencí. Z asijských zemí, např. ze Saudské Arábie nebo z Kuvajtu, je odesláno významné množství remitencí, naopak přijato není téměř nic. Německo je příkladem země, kde odeslané a přijaté remittance jsou téměř na stejné úrovni.



Obrázek 39 Objem odeslaných a přijatých remitencí ve světě (2012, mld. USD)

Zdroj: The World Bank 2013, vlastní zpracování

Po přepočtení rozdílu odeslaných a přijatých remitencí k výši HDP je patrné, že nejvíce ohroženou ekonomikou z hlediska odeslaných remitencí je Lucembursko. V roce 2012 tvořily odeslané remittance téměř 20 % HDP, v opačném směru do této země přiteklo pouze necelé 3 % HDP (The World Bank, 2013). Další evropskou zemí, na kterou má migrace z hlediska odeslaných remitencí negativní vliv, je Švýcarsko. Rozdíl odeslaných a přijatých remitencí zde tvořil v roce 2012 4 % HDP.



Obrázek 40 Podíl rozdílu odeslaných a přijatých remitencí na HDP (2012, %)

Zdroj: The World Bank 2013, vlastní výpočty

V souvislosti s odhadem remitencí je nutno podotknout, že shromažďování přesných údajů týkajících se remitencí je nesmírně obtížný úkol a skutečné toky remitencí jsou obvykle podhodnoceny. Jedním z důvodů je to, že v oficiálních statistikách nejsou zachyceny finanční převody mimo bankovní systém (Schiopu, Siegfried, 2006). Neformální toky představují v některých případech poměrně vysoké částky a jejich výše se podle odhadů také velmi liší.

Závěr

Většina studií zabývajících se remitencemi považuje mezinárodní migraci a posílané remitence jako nástroj rozvoje pro ekonomicky chudší země. Na druhou stranu ovšem bývá opomíjen fakt, že pro země, kde podíl remitencí na hrubém domácím produktu je vysoký, dochází k velkému odlivu vytvořeného bohatství do zahraničí. Představená problematika patří do řídce diskutovaných oblastí na pomezí demografie a ekonomické statistiky. Význam pro tuzemský výzkum umocňuje to, že Česká republika dnes patří mezi země imigrační, kde podíl cizinců na celkové populaci je zhruba 5 % a důvody pro imigraci do Česka jsou převážně pracovní. Málo diskutovanou oblast tvoří to, že někteří imigranti vyhledávají práci v cizině zejména proto, aby finančně zabezpečili svou rodinu v mateřské zemi. Vedle zřejmého pozitivního vlivu migrace spočívajícího v přílivu kvalifikované i nekvalifikované a potřebné pracovní síly tak má migrace i mírně negativní vliv na českou ekonomiku, právě pokud jde množství odeslaných remitencí. Obavy z mizivého pozitivního vlivu nebo dokonce negativního vlivu migrace se nepotvrzují. V tomto článku bylo ukázáno, že velká část odeslaných prostředků se vrátí v podobě přijatých remitencí od občanů České republiky, kteří pracují v zahraničí. Lze konstatovat, že v současné době migrace významně neovlivňuje situaci v České republice a do budoucna sama o sobě zřejmě důsledky stárnutí populace nezmírní. Česká republika by měla nadále podporovat všechny formy legální pracovní migrace, pokud chce zmírnit důsledky stárnutí populace a úbytek obyvatel v produktivním věku. Měla by usilovat o stabilní a kvalifikovanou pracovní sílu. Otázky, které je třeba v souvislosti se stárnutím řešit, by tedy měly zahrnovat nejen věkovou hranici pro odchod důchod, úroveň a povahu zdravotních dávek pro seniory, ale také účast na trhu práce a politiku a programy týkající se mezinárodní migrace, zejména náhradové migrace a integrace velkého počtu imigrantů a jejich potomků. Nejedná se jen o počet imigrantů, ale také o jejich strukturu. Žádoucí jsou imigranti, kteří se zde natrvalo usadí a založí rodinu. Společně s prodlužováním věku odchodu do důchodu by tak migrace mohla částečně zmírnit dopady na ekonomiku České republiky způsobenou stárnutím populace.

Literatura

- ČSÚ, 2013a. *Cizinci v ČR 2013*. Český statistický úřad. Praha [online]. Dostupný z WWW: <<http://www.czso.cz/csu/2013edicniplan.nsf/p/1414-13>> [cit. 18.1.2014].
- ČSÚ, 2013b. *HDP Výrobní metoda*. Databáze ročních národních účtů. Český statistický úřad. Praha [online]. Dostupný z WWW: <http://apl.czso.cz/pll/rocnka/rocnkavyber.makroek_prod> [cit. 10.1.2014].
- EUROSTAT, 1996. *European System of Accounts – ESA 1995*. Luxembourg, Office for Official Publications of the European Communities.
- IMF 2009. *International transactions in remittances: guide for compilers and users*. Washington D.C.: International Monetary Fund.
- LEONTIYEVA Y., TOLLAROVÁ B., 2011. „Šetření cizinců o jejich příjmech, výdajích a remitencích. Závěrečná zpráva z výzkumu“. Praha: Sociologický ústav AV ČR, v.v.i. [online]. Dostupný z WWW: <<http://www.soc.cas.cz/articles/cz/5/6127/Leontiyeva-Y.-B.-Tollarova-2011.-8222-Results-from-a-Survey-of-Foreigners-8217-Incomes-Expenditures-and-Remittances.-Main-Findings-Concerning-Remittances-8220-.-Praha-Sociologicky-ustav-AV-CR-v.v.i..html>> [cit. 2.12.2013].
- MŠMT, 2009. *ÚIV*. Ministerstvo školství, mládeže a tělovýchovy. Praha [online]. Dostupný z WWW: <<http://www.msmt.cz/vzdelavani/socialni-programy/dokumenty-6>> [cit. 13.9.2013]
- MPSV, 2011a. *Důchodová reforma – Přílohy*. Tisková zpráva Ministerstva práce a sociálních věcí. [online]. Dostupný z WWW: <http://www.mpsv.cz/files/clanky/10407/22022011_p1.pdf> [cit. 10.1.2014]
- MPSV, 2011b. *Souhrnná informace za rok 2011 o aktivitách realizovaných příslušnými resorty v oblasti potírání nelegálního zaměstnávání cizinců* [online]. Ministerstvo práce a sociálních věcí. Praha [online]. Dostupný z WWW: <http://www.mpsv.cz/files/clanky/13355/potirani_cerne_prace_2011.pdf> [cit. 18.1.2014].

MZV, 2013. *Zákon č. 326/1999 sb. o pobytu cizinců na území ČR a o změně některých zákonů*. Ministerstvo zahraničních věcí. Praha [online]. Dostupný z WWW: <<http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=5978>> [cit. 10.1.2014].

ONDRUŠ V., 2009. *National Accounts and economic migration – Remittances in the Czech Republic*. 17th International Input-output Conference. Sao Paulo, Brazil 2009 [online]. Dostupný z WWW: <http://www.iioa.org/conferences/17th/papers/1068331043_090529_132336_PAPER_ONDRUS.PDF> [cit. 2.9.2013].

PARK S., 2007. *Does Immigration Benefit a Regional Economy With An Aging Population? Simulation Results from the Chicago CGE Model*. Department of Economics. University of Illinois at Urbana-Champaign [online]. Dostupný z WWW: <<http://www.law.berkeley.edu/files/Parkpaperfinal.pdf>> [cit. 10.9.2013]

SCHIOPU I., SIEGFRIED N., 2006. *Determinants of workers' remittances. Evidence from the European neighbouring region*. European Central Bank [online]. Dostupný z WWW: <<http://www.suomenpankki.fi/pdf/128539.pdf>> [cit. 18.1.2014].

THE WORLD BANK, 2013. *Migration and Remittances. Annual remittances data*. The World Bank group [online]. Dostupný z WWW: <<http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTDECPROSPECTS/0,,contentMDK:22759429~pagePK:64165401~piPK:64165026~theSitePK:476883,00.html>> [cit. 10.1.2014].

JEL Classification: F22, F24, J61

Summary

The labour migration in the Czech Republic

This paper deals with the issue of migration in connection with aging of Czech population. Because the proportion of older people is increasing and in the same way as the proportion of productive population is declining, migration for work is one of the factors to ensure stable level of the workforce. The issue of migration is closely related to phenomenon of remittances as outflow resources abroad. This article introduces the number and structure of migrants in the Czech Republic and the amount of remittances sent abroad as well as the population of the Czech Republic working abroad and the amount of arriving remittances. This analysis allows decision that migration and remittances don't significantly affect the Czech economy, positively or negatively. Finally, I also made international comparison with countries in which migration in terms of remittances sent presents a significant issue.

Key words: aging of population, labour market, labour migration, remittances.

Modelování demografických časových řad pro potřeby demografických projekcí

Ondřej Šimpach

ondrej.simpach@vse.cz

Doktorand oboru statistika

Školitel: doc. Ing. Jitka Langhamrová, CSc., (langhamj@vse.cz)

Abstrakt: Studie pojednává o hlavních rozdílech mezi deterministickými a stochastickými přístupy při tvorbě demografických projekcí. V minulosti se využívaly výhradně projekce deterministické, které byly sice výpočetně jednoduché, ale v současné době již nestačí na kvalitní vysvětlení odhadovaného budoucího trendu. Vzhledem k tomu, že čistě stochastický přístup je ale zase poměrně výpočetně náročný a velké nároky jsou kladeny i na množství a kvalitu vstupních dat, bude představen přístup individuálního modelování časových řad Boxovou-Jenkinsonovou metodologií na vybraný demografický ukazatel – věkově specifické míry úmrtnosti mužů a žen, neboť odhadovaný budoucí vývoj úmrtnosti analyzované populace je jedna z klíčových vstupních informací pro každou populační projekci. K dispozici bude datová základna Českého statistického úřadu od roku 1920 do roku 2012. S využitím zmíněné metodologie budou konstruovány individuální modely ARIMA (s konstantou nebo bez ní) pro osoby 0–100+ leté a dále individuální modely náhodné procházky s konstantou, zvlášť pro mužské a zvlášť pro ženské pohlaví. Na základě těchto modelů budou konstruovány předpovědi ukazatele věkově specifických měr úmrtnosti na období let 2013 až 2050. Získané předpovědi budou spolu konfrontovány a v závislosti na poskytnutých výsledcích a jejich výpočetní náročnosti bude stanoveno doporučení o jejich vhodnosti, případně o jejich dalším zefektivnění. Nalezením univerzálního modelu pro všechny věkové skupiny v konkrétní populaci by došlo ke zrychlení výpočetního procesu a možnosti častější rekalkulace demografických projekcí.

Klíčová slova: věkově specifické míry úmrtnosti, časové řady, ARIMA, náhodná procházka, projekce

Úvod

Demografické projekce možného budoucího vývoje populace jsou základním informačním kanálem, který slouží k poskytnutí klíčových informací o potenciálním vývoji úmrtnosti, porodnosti, imigrace a emigrace, popřípadě dalších demografických statistik. Každá projekce je založena na předpokladech, které do jisté míry mohou, nebo nemusí nastat. Obecně z přístupů o vývoji poklesu úmrtnosti populace, o kterých pojednával již Halley (1693), a byl dále rozveden a zpracován Malthusem (1798a, 1798b), můžeme tvrdit, že jsou projekce *deterministické*, tj. založené na předem stanovených předpokladech. Předpoklady byly zpočátku stanovovány expertními úsudky, později již podpůrnými statistickými metodami, které se dále rozvíjely a zdokonalovaly. Mezi tyto podpůrné metody bývají zahrnovány lineární či vícenásobné regrese, do expertních úsudků patří odhady založené na doporučeních odborníky z oblasti sociologie, politologie, medicíny a práva. Kombinací těchto dvou přístupů můžeme dospět k očekávaným scénářům, které dále slouží jako vstupní informace k procesu samotného projekčního výpočtu. Pro deterministické typy demografických projekcí se nejčastěji využívá přístup komponentní metody, která se dále nechá obohacovat a rozvíjet o prvky z teorie pravděpodobnosti. Komponentní metoda je velmi starý algoritmus, nicméně pro svou užitečnost a zejména jednoduchost je oblíbeným projekčním nástrojem dodnes.

Zcela odlišný přístup demografických projekcí, o kterých se zmiňuje např. Bell, Monsell (1991), Lee, Carter (1992) nebo Lee, Tuljapurkar (1994), jsou projekce *stochastické*. Tyto projekce jsou založeny na hledání hlavních komponent, vysvětlující trend, který je zakomponován ve vývoji časových řad věkově specifických demografických měr. Velkou roli na výsledky hraje délka časových řad (viz např. Coale, Kisker (1986), nebo srovnání výsledků více populací ze studie Booth et al. (2005)). Existují země, které mají své statistiky velmi podrobné za velmi dlouhé časové období, ale jsou také země, které takto podrobné statistiky nemají, nevedly je, nebo z důvodu politických, ekonomických či společenských událostí došlo k narušení kontinuity v pořizování těchto údajů. Hlavními důvody jsou územně-politické transformace, zániky a vzniky nových států a bohužel také občanské a světové války. Česká republika na tom po datové stránce není ani nejhůře, ale ani nejlépe. Stochastickými projekcemi se mohli zabývat

např. Lee, Carter (1992), kteří byli schopni analyzovat trendy ve vývojích časových řad za Spojené státy americké, neboť podrobné statistiky vykazují údaje již od deklarace americké nezávislosti. Velká Británie a země britského společenství jsou po datové stránce také výborně vybaveny, neboť sofistikované projekce pro Velkou Británii zpracoval např. Murphy (1995), v Kanadě se stochastickým modelováním úmrtnosti zabýval také Keyfitz (1991), a v současné době velmi diskutovaná témata zpracovává např. Mélard, Pasteels (2000), Lundström, Qvist (2004), Hyndman, Shang, (2009) nebo Erbas et al. (2012), kteří vyvíjí sofistikovaný software a zdokonalují jednotlivé přístupy pro stochastické demografické prognózování. Již první zmínky o automatizovaném prognózování uvedli Hyndman et al. (2002) v případě australské populace, později se Hyndman (2012) začal zabývat zobecňováním a experimenty z různých typů časových řad a různými typy populací. Inspirace byla také čerpána od Orda a Loweho (1996).

Za Českou republiku není k dispozici tak kvalitní datová základna jako za Spojené státy nebo za Velkou Británii a země britského společenství. V současné době jsme schopni připravit datovou matici vybraných demografických ukazatelů podle věku a pohlaví maximálně od roku 1920 do současnosti, některé jen od roku 1950, neboť druhá světová válka a zřízený Protektorát Čechy a Morava sběr a správnost prezentovaných dat ve válečném období zkomplikovaly. (Použitá datová základna pro tuto studii pochází z databáze Českého statistického úřadu (ČSÚ)). Nicméně existují i země, které jsou mladší než Česká republika, a jejich přepočty původních dat z původní populace (např. po transformaci) se nezdařil. Pak tyto země mají datovou základnu velmi krátkou a přístupy pro modelování časových řad zde použít nejdou. (Pomiňme nyní situace např. některých afrických zemí, které statistiky buď prakticky nevedou vůbec, nebo jsou zcela nevypovídající a nekorektní. Pak nelze použít vůbec žádný přístup).

Přístupy založené na *deterministickém* základu se v podmínkách České republiky používají dodnes. Můžeme zmínit např. studii Koschina (2005), Fialy (2006) nebo Koschina at al. (2007). Zejména pak studie Fialy (2006) poskytuje informaci, že $\ln(m_x)$, kde m_x je věkově specifická míra úmrtnosti, je pro každý věk x přibližně lineární. Z toho lze usoudit, že podaří-li se nám nalézt vhodnou úrovněovou konstantu (b_0) a sklon (b_1) regresní přímky, můžeme snadno provést lineární extrapolaci do budoucna jako

$$\ln(m_{x,t}) = b_0 + b_1 t, \quad (1)$$

kde t je konkrétní rok a naprojektovat tak očekávaný vývoj budoucích specifických měr. Tento přístup je jednoduchý, nicméně vzhledem k tomu, že data v podmínkách České republiky jsou poměrně variabilní a charakter časových řad jednotlivých věkově specifických měr je rozkolísaný, rezidua z těchto regresí by byla silně autokorelovaná (viz Aitken, 1935, Arlt, Arltová, 2007) a sklon regresní přímky, (zejména pak v nejnižších a nejvyšších věkových skupinách, kde je nejvyšší variabilita v datech), by nemusel být optimální k dosažení žádoucích výsledků. Budeme tvrdit, že tento přístup není v dnešní době již postačující a je třeba přejít na jiný. Čistě *stochastický* přístup, založený zejména na metodě Lee-Cartera (Lee, Carter, 1992) a kointegrační analýze (Engle, Granger, 1987), se již v podmínkách České republiky také využil, např. ve studii Arltové et al. (2013). Tento přístup je vysoce sofistikovaný, ale také poměrně výpočetně náročný a komplikovaný. Víme, že mezi stochastické modelování časových řad patří i jednorozměrné modely, založené na metodologickém přístupu Boxe, Jenkinse, (1970). Rodí se tedy otázka, zdali jednoduchý deterministický základ, který byl zejména pro potřeby České republiky hojně využíván doposud, částečně nepřetransformovat ve stochastický přístup? (viz např. srovnání různých přístupů předpovídání demografických časových řad Bellem, 1997). Vybraná data, zejména pak počty zemřelých x -letých podle pohlaví (M_x), počty obyvatel k 1.7. daného roku podle pohlaví x -letých ($\bar{S}_x$) a nakonec i počty živě narozených daného roku podle pohlaví (N_x^v), máme k dispozici od roku 1920 do roku 2012. Ze vztahu

$$\ln(m_{x,t}) = \ln\left(\frac{M_{x,t}}{\bar{S}_{x,t}}\right), \quad (2)$$

kde $x = 0, 1, 2, \dots, 100+$, získáváme 101 časových řad mužských logaritmů věkově specifických měr úmrtnosti a 101 časových řad ženských logaritmů věkově specifických měr úmrtnosti, obojí o četnosti 92 ročních pozorování. (Logaritmy jsou použity z toho důvodu, aby byl lépe patrný průběh jednotlivých časových řad, neboť samotné specifické míry úmrtnosti jsou mezi věky 5–40 let velmi blízké nule).

Podařilo-li by se nalézt vhodné modely pro tyto časové řady, které vystihnou trend, který se ve zmíněných časových řadách nachází, pak by mohly být vypočteny předpovědi na rozumně dlouhé období. Tyto předpovědi věkově specifických měr úmrtnosti by pak byly základem pro výpočet úmrtnostních tabulek a všech odvozených ukazatelů z nich plynoucích, a dále pochopitelně jako vstupní scénář pro výpočet kompletní populační projekce.

Cílem předkládané studie je pokusit se odhadnout 101 modelů pro mužské a 101 modelů pro ženské pohlaví, na základě modelů $ARIMA(p,d,q)$ s konstantou nebo bez ní, založených na metodologickém přístupu stochastického modelování časových řad, a dále 101 modelů pro mužské a 101 modelů pro ženské pohlaví, na základě modelů náhodné procházky s konstantou. Na základě těchto 4 x 101 modelů bude vypočteno 4 x 101 předpovědí logaritmů věkově specifických měr úmrtnosti na období let 2013–2050, což je 38 pozorování pro každý věk x a každé pohlaví. Tyto předpovědi budou spolu konfrontovány a v závislosti na poskytnutých výsledcích a jejich výpočetní náročnosti bude stanoveno doporučení o jejich vhodnosti, případně o jejich dalším zefektivnění. Výpočet modelů $ARIMA$ bude prováděna v systému Statgraphics Centurion XVI., výpočet modelů náhodných procházek s konstantou tamtéž. Prezentace výsledků logaritmů věkově specifických měr bude prováděna v grafickém prostředí systému RStudio verze 0.98.490, neboť se bude jednat o rozsáhlé datové matice.

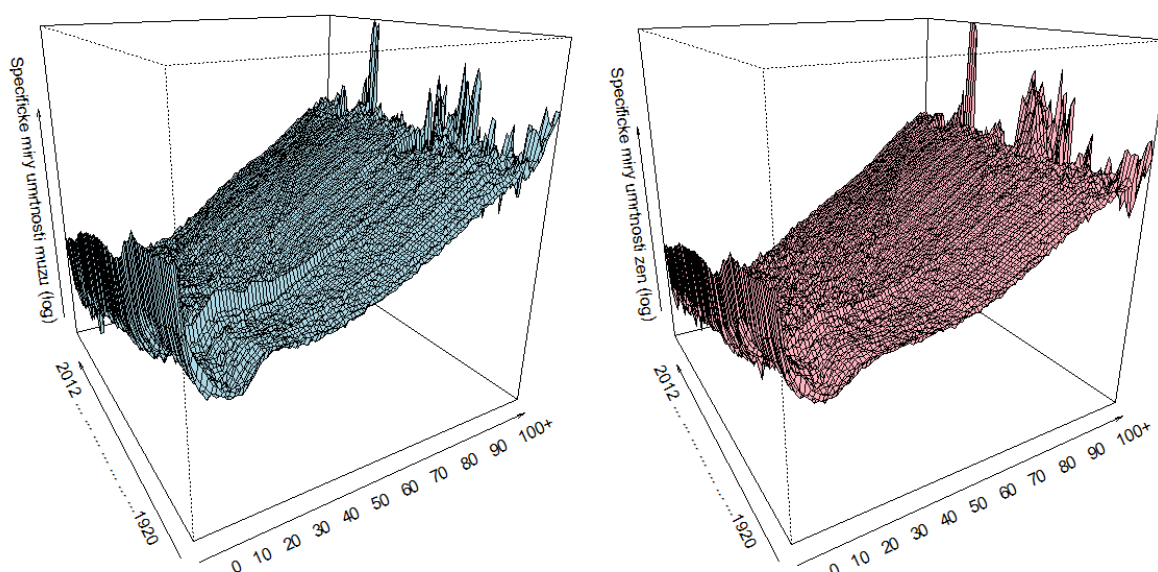
Metodika

Data, která budou analyzována v této studii, můžeme chápat jako početnou skupinu časových řad. Vzhledem k tomu, že věkově specifické míry úmrtnosti začínají věkem 0 dokončených let a končí věkem 100+, což jsou osoby stoleté a všechny starší, uvažujeme 101 časových řad pro mužské a dalších 101 časových řad pro ženské pohlaví. Každá specifická míra úmrtnosti, pro každý věk a každé pohlaví, má svůj trend, který se vytvářel v průběhu vývoje v letech 1920–2012. Každá z těchto v sumaci 202 časových řad má svých 92 pozorování, která budou modelována Boxovou-Jenkinsovou metodologií (Box, Jenkins 1970). Lineární modely Boxovy-Jenkinsovy metodologie, představené např. Gardnerem Jr., McKenzie (1985) nebo Arltem (1999) vycházejí z tzv. lineárního procesu. Zmíněným metodologickým přístupem budou odhadnuty parametry modelů $ARIMA(p,d,q)$ s konstantou c nebo bez ní, přičemž p značí počet zahrnutých členů AR (auto-regresních), d je počet diferencí a q značí počet zahrnutých členů MA (klouzavých průměrů – moving averages). Parametry byly vybrány tak, byly respektovány požadavky na hodnoty parametrů a dále diagnostické testy modelu, zejména pak Boxův-Pierceův test autokorelace reziduí (Box, Pierce, 1970, nebo Jarque, Bera, 1980), kde bylo vyžadováno nezamítnutí testované nulové hypotézy o sériové nekorelovanosti reziduí. Ve druhé fázi pak byly vypočteny modely náhodné procházky s konstantou.

Provedené experimenty

Pohledme nyní na empirická data logaritmů věkově specifických měr úmrtnosti mužů a žen s pomocí prostorového X - Y - Z grafu, kde na ose X odečítáme věk x -letých osob, na ose Y odečítáme čas t a na ose Z odečítáme příslušné hodnoty logaritmů těchto měr. Míry jsou zobrazeny na obrázku 1 (vlevo pro muže, vpravo pro ženy). Zejména u mužské populace je výrazně viditelné období druhé světové války a dále obě pohlaví zaznamenávají poměrně vysokou variabilitu v empirických datech u nejvyšších věkových skupin.

Na základě přístupu stochastického modelování časových řad, zmíněném v kapitole Metodika, byly odhadnuty modely $ARIMA(p,d,q)$ s konstantou c nebo bez ní, a sumarizace těchto výsledků je uvedena pro mužskou populaci v tabulce 1, respektive pro ženskou populaci v tabulce 2.



Obrázek 1: Logaritmy věkově specifických měr úmrtnosti pro muže (vlevo) a pro ženy (vpravo) v České republice v letech 1920–2012

Zdroj dat: ČSÚ, vlastní výpočet a konstrukce.

Tabulka 1: Modely ARIMA (p,d,q) s konstantou c nebo bez ní pro mužské logaritmy věkově specifických měr úmrtnosti osob 0–100+ v České republice.

Zdroj: vlastní výpočet

Věk	Model	Věk	Model	Věk	Model	Věk	Model	Věk	Model	Věk	Model	Věk	Model	Věk	Model
0	(0,2,1) c	13	(0,1,1) c	26	(0,1,1) c	39	(0,1,1) c	52	(0,1,1) c	65	(0,1,1) c	78	(0,1,1) c	91	(0,1,2) c
1	(0,1,1) c	14	(0,1,1) c	27	(0,1,1) c	40	(2,2,1)	53	(0,1,1) c	66	(0,1,1) c	79	(1,1,1) c	92	(2,1,0) c
2	(2,2,1)	15	(0,1,1) c	28	(0,1,1) c	41	(0,1,1) c	54	(0,1,1) c	67	(0,1,1) c	80	(0,1,1) c	93	(0,1,1) c
3	(0,1,1) c	16	(0,1,1) c	29	(0,1,1) c	42	(0,1,1) c	55	(0,1,1) c	68	(0,1,1) c	81	(0,1,1) c	94	(0,1,2) c
4	(0,1,1) c	17	(0,1,1) c	30	(0,1,1) c	43	(0,1,1) c	56	(2,1,0) c	69	(0,1,1) c	82	(0,1,1) c	95	(1,1,1) c
5	(0,1,1) c	18	(0,1,1) c	31	(2,2,1)	44	(0,1,1) c	57	(0,1,1) c	70	(0,1,1) c	83	(1,1,1) c	96	(0,1,2) c
6	(0,1,1) c	19	(0,1,1) c	32	(0,1,1) c	45	(0,1,1) c	58	(0,1,1) c	71	(0,1,1) c	84	(0,1,1) c	97	(0,1,1) c
7	(0,1,1) c	20	(0,1,2) c	33	(0,1,1) c	46	(0,1,1) c	59	(0,1,1) c	72	(0,1,1) c	85	(0,1,1) c	98	(0,1,1) c
8	(0,1,1) c	21	(0,1,1) c	34	(0,1,1) c	47	(2,1,0) c	60	(0,1,1) c	73	(1,1,0) c	86	(0,1,1) c	99	(0,1,2) c
9	(0,1,1) c	22	(0,1,1) c	35	(0,1,1) c	48	(1,1,0) c	61	(0,1,1) c	74	(1,1,0) c	87	(0,1,1) c	100+	(0,1,1) c
10	(0,1,1) c	23	(0,1,1) c	36	(0,1,1) c	49	(0,1,1) c	62	(0,1,1) c	75	(2,1,0) c	88	(0,1,1) c		
11	(0,1,1) c	24	(0,1,1) c	37	(0,1,1) c	50	(2,1,0) c	63	(0,1,1) c	76	(0,1,1) c	89	(0,1,2) c		
12	(0,1,1) c	25	(2,1,0) c	38	(0,1,1) c	51	(1,1,0) c	64	(0,1,1) c	77	(0,1,1) c	90	(2,1,0) c		

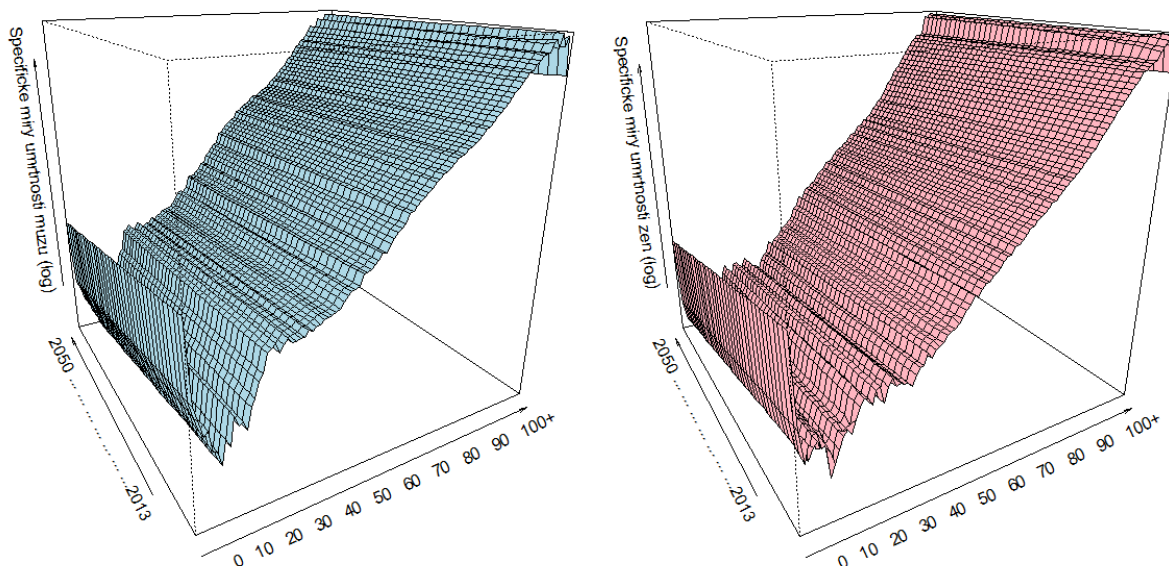
Tabulka 2: Modely ARIMA (p,d,q) s konstantou c nebo bez ní pro ženské logaritmy věkově specifických měr úmrtnosti osob 0–100+ v České republice

Zdroj: vlastní výpočet

Věk	Model	Věk	Model	Věk	Model	Věk	Model	Věk	Model	Věk	Model	Věk	Model	Věk	Model
0	(0,2,1) c	13	(0,1,1) c	26	(0,1,1) c	39	(0,1,1) c	52	(0,1,1) c	65	(1,1,0) c	78	(2,1,2) c	91	(1,1,1) c
1	(0,1,1) c	14	(1,1,1) c	27	(0,1,1) c	40	(2,1,0) c	53	(0,1,1) c	66	(1,1,0) c	79	(0,1,1) c	92	(2,1,0) c
2	(1,1,1) c	15	(0,1,1) c	28	(0,1,1) c	41	(1,1,0) c	54	(0,1,1) c	67	(0,1,1) c	80	(0,1,1) c	93	(0,1,1) c
3	(2,1,0) c	16	(2,1,0) c	29	(0,1,1) c	42	(0,1,1) c	55	(0,1,1) c	68	(0,1,1) c	81	(0,1,1) c	94	(2,1,1) c
4	(2,1,0) c	17	(0,1,1) c	30	(0,1,1) c	43	(0,1,1) c	56	(0,1,1) c	69	(0,1,1) c	82	(0,1,1) c	95	(1,0,0) c
5	(0,1,1) c	18	(0,1,2) c	31	(0,1,1) c	44	(0,1,1) c	57	(0,1,1) c	70	(0,1,1) c	83	(0,1,1) c	96	(1,0,0) c
6	(1,1,1) c	19	(0,1,1) c	32	(0,1,1) c	45	(0,1,1) c	58	(2,1,0) c	71	(1,1,0) c	84	(0,1,1) c	97	(0,0,2) c
7	(0,1,1) c	20	(0,1,1) c	33	(0,1,1) c	46	(0,1,1) c	59	(1,1,1) c	72	(2,1,0) c	85	(0,1,1) c	98	(0,0,2) c
8	(2,1,0) c	21	(0,1,1) c	34	(0,1,1) c	47	(0,1,1) c	60	(0,1,1) c	73	(1,1,0) c	86	(0,1,1) c	99	(1,1,0) c
9	(2,1,0) c	22	(2,1,0) c	35	(2,1,0) c	48	(0,1,1) c	61	(0,1,2) c	74	(2,1,0) c	87	(0,1,1) c	100	(0,1,1) c
10	(0,1,1) c	23	(0,1,1) c	36	(0,1,1) c	49	(1,1,1) c	62	(1,1,0) c	75	(2,1,0) c	88	(0,1,1) c		
11	(0,1,1) c	24	(1,1,0) c	37	(0,1,1) c	50	(0,1,1) c	63	(1,1,0) c	76	(0,1,1) c	89	(1,1,1) c		
12	(1,1,1) c	25	(0,1,1) c	38	(0,1,1) c	51	(0,1,1) c	64	(0,1,1) c	77	(0,1,1) c	90	(2,1,0) c		

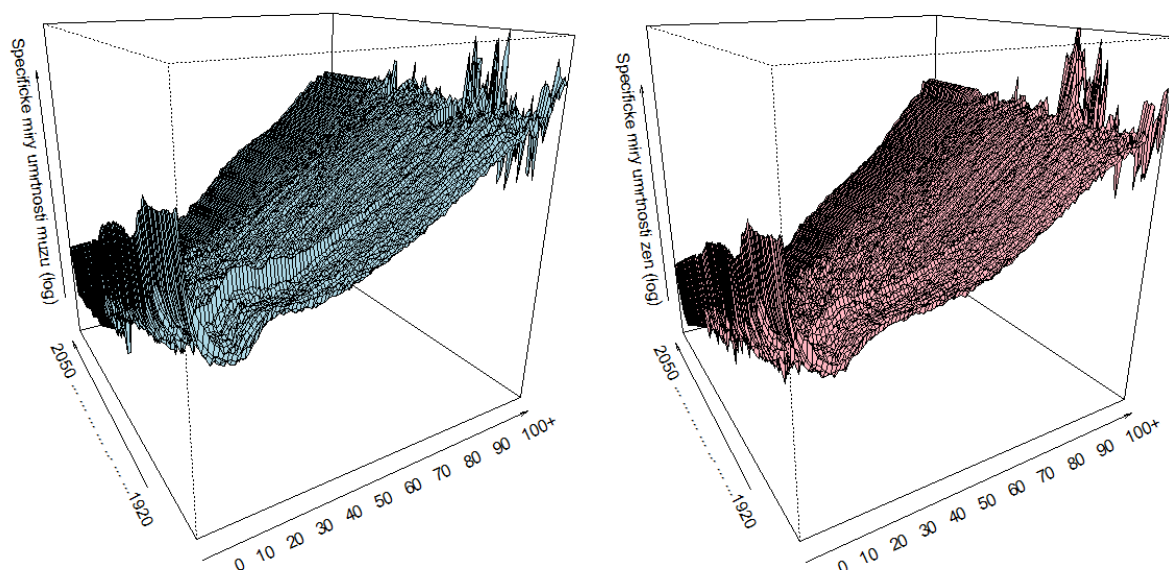
Z uvedených modelů, u kterých byl kladen důraz na projití diagnostickými testy, byly vypočteny předpovědi logaritmu mužských a ženských věkově specifických měr úmrtnosti na období let 2013 až 2050.

Tyto předpovědi byly zobrazeny X-Y-Z prostorovým grafem, a prezentuje je obrázek 2, (vlevo pro mužské a vpravo pro ženské pohlaví). Z uvedeného výstupu můžeme usoudit, že obecný předpoklad budoucího poklesu hodnot věkově specifických měr úmrtnosti platí, neboť predikované časové řady pro jednotlivé roky života mají tendenci viditelně do budoucna klesat. Tento viditelný pokles bude patrný v obrázku 3, kde bylo provedeno napojení těchto predikovaných věkově specifických měr na empirická data.



Obrázek 2: Predikce logaritmů věkově specifických měr úmrtnosti pro muže (vlevo) a pro ženy (vpravo) modely ARIMA pro Českou republiku v letech 2013–2050

Zdroj: vlastní výpočet.

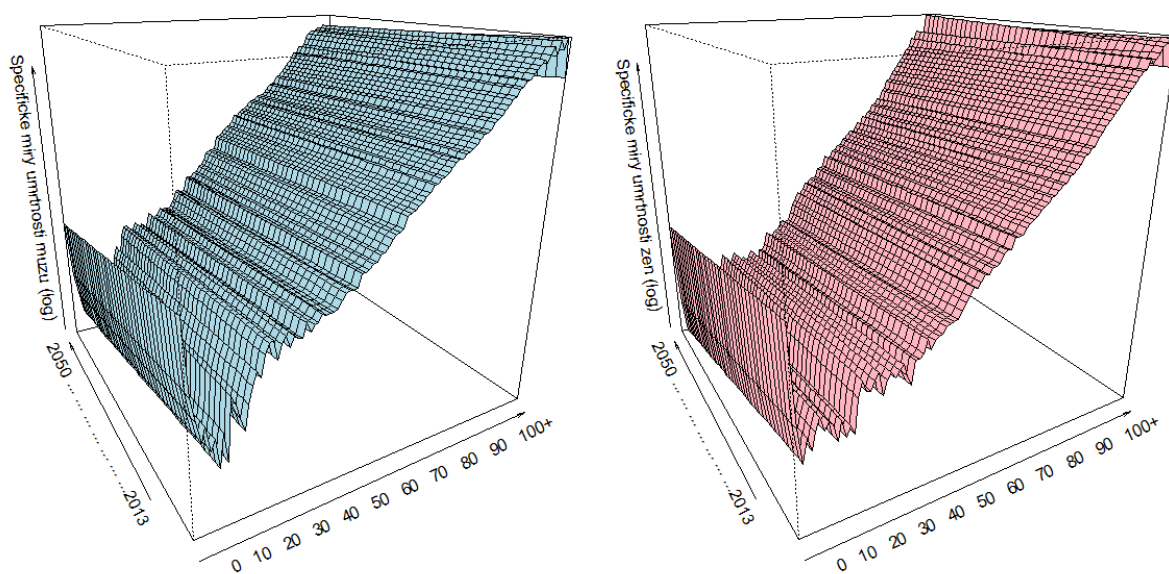


Obrázek 3: Vývoj logaritmů věkově specifických měr úmrtnosti pro muže (vlevo) a pro ženy (vpravo) v České republice v letech 1920–2012 s připojenými predikovanými logaritmy těchto měr pro období 2013–2050, vypočtenými modely ARIMA

Zdroj: ČSÚ, vlastní výpočet a konstrukce.

Na obrázku 3 pozorujeme poměrně plynulý přechod z empirických dat na predikovaná data, jak u mužské, tak u ženské populace. Je pochopitelné, že v nejvyšších věkových skupinách se již nevyskytují žádné extrémní hodnoty, neboť byly modelem vyhlazeny. Ani vysoká variabilita v empirických datech však nepoškodila model natolik, aby v nejnižších i nejvyšších věkových skupinách nebyl významný a vykazoval nereprezentativní výsledky. Obecná očekávání o poklesu věkově specifických měr úmrtnosti v čase jsou zachována, do budoucna jejich hodnoty klesají.

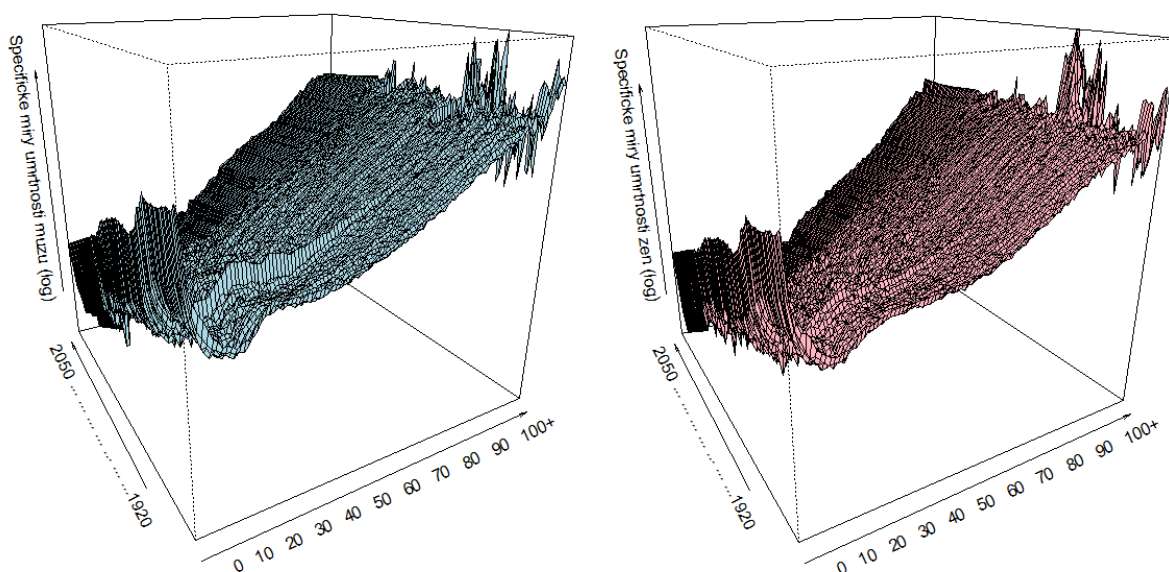
Jako druhý typ modelu, který byl zvolen pro vyrovnání empirických dat a následný výpočet předpovědí, je model náhodné procházky s konstantou. Pro 101 mužských a 101 ženských časových řad logaritmů věkově specifických měr úmrtnosti byly vypočteny konstanty a na jejich základě zkonstruovány předpovědi. Ve všech případech kromě věku 98 dokončených let u ženského pohlaví tyto modely fungují a poskytují výsledky dle obecných očekávání poklesu úmrtnostních měr ve všech věkových skupinách do budoucna. Z důvodu velmi vysoké variability v časové řadě pro ženy dokončeného věku 98 let nastala situace, že se očekávaný budoucí vývoj vychýlil do opačného směru (nepatrně narůstá). Tuto nepříznivou situaci bychom museli vyřešit například korekcí extrémních hodnot v časové řadě. V žádném jiném případě tato situace nenastává a předpovědi mužských (vlevo) a ženských (vpravo) logaritmů věkově specifických měr úmrtnosti můžeme vidět na obrázku 4.



Obrázek 4: Predikce logaritmů věkově specifických měr úmrtnosti mužů (vlevo) a žen (vpravo) modely náhodné procházky s konstantou pro Českou republiku v letech 2013–2050

Zdroj: vlastní výpočet.

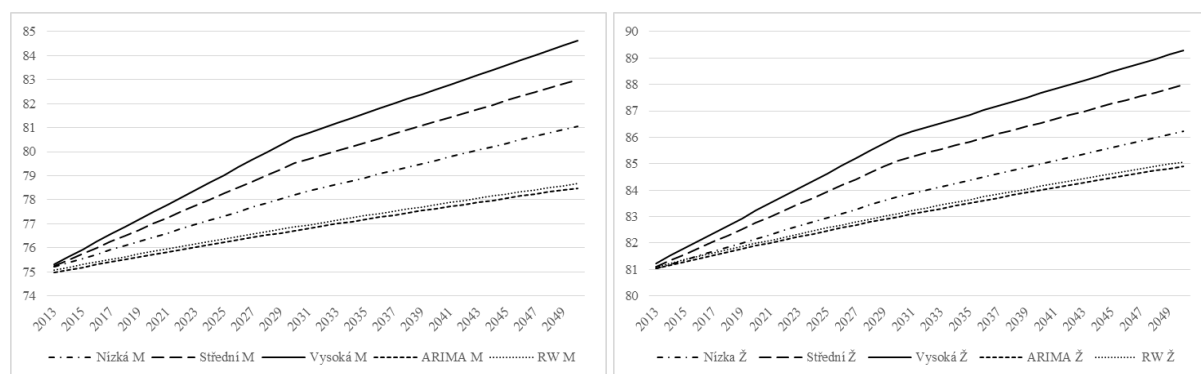
Stejně jako v případě optimalizovaných modelů ARIMA bude výhodné, když v jednom grafu napojíme empirické hodnoty logaritmů věkově specifických měr na vypočtené predikované hodnoty. Celý vývoj časové řady těchto měr je patrný v obrázku 5, kde jako v předchozích případech vidíme v levé části mužské, v pravé ženské pohlaví. V napojení predikovaných hodnot není patrný žádný extrémní přechod, můžeme tvrdit, že je velmi plynulé. Jediný fakt, se kterým můžeme být nespokojeni je ona skutečnost mírně narůstajících hodnot logaritmů věkově specifických měr úmrtnosti žen ve věku 98 dokončených let (viz horní část pravého grafu v obrázku 5).



Obrázek 5: Vývoj logaritmu věkově specifických měr úmrtnosti mužů (vlevo) a žen (vpravo) v České republice v letech 1920–2012 s připojenými predikovanými logaritmy těchto měr pro období 2013–2050, vypočtenými modely náhodné procházky s konstantou

Zdroj: ČSÚ, vlastní výpočet a konstrukce.

Nyní ve fázi, kdy máme vypočtené všechny předpovědi, (jak modely ARIMA tak náhodné procházky), můžeme zkusit srovnat scénář, který vyplyne z naší vypočtených hodnot, se scénářem, se kterým kalkuluje např. Český statistický úřad. Populační projekce, vypočtená Českým statistickým úřadem v roce 2013 (ČSÚ, 2013b), je stejně jako všechny předchozí, počítána komponentní metodou čili deterministicky. Zmíněná populační projekce počítá se snižováním úmrtnostních měr, ale možná příliš výrazně. V případě, že provedeme odlogaritmování našich predikovaných věkově specifických měr úmrtnosti, vypočtených na základě stochastického modelování jednotlivých časových řad, můžeme s využitím algoritmu výpočtu klasických úmrtnostních tabulek (ČSÚ, 2013a) dospět k odvozeným ukazatelům, zejména pak ke střední délce života – (naději dožití) osoby 0leté. Nízkou, střední a vysokou variantu populační projekce ČSÚ tedy nyní srovnáme s našimi projekcemi ve společném grafu. Jak v případě mužů, tak v případě žen, jsou předpovědi modely ARIMA i náhodné procházky s konstantou pod úrovní nízké varianty ČSÚ (viz obrázek 6, kde v levé části je zobrazeno mužské, v pravé části ženské pohlaví). Může to být způsobeno skutečností, že ČSÚ může mít své odhady nadhodnoceny právě z důvodu použití deterministického modelu, který trend nebere v úvahu tak, jako stochastický model.



Obrázek 6: Střední délka života 0letých osob pro muže (vlevo) a pro ženy (vpravo), vyplývající z populační projekce České republiky, vypočtené statistickým úřadem a dále vyplývající z předpovědí modely ARIMA a náhodné procházky s konstantou

Zdroj: ČSÚ, vlastní výpočet a konstrukce.

Závěr

Studie ukázala pokus o alternativní přístup k odhadu budoucího vývoje logaritmu věkově specifických měr úmrtnosti, které jsou nezbytnou součástí pro výpočty demografických populačních projekcí. Přístup je založen na stochastickém modelování časových řad, přičemž časové řady byly analyzovány individuálně, bez ohledu na případnou vzájemnou provázanost. (Vzájemnou provázanost však nerespektuje ani deterministický model, který se v populačních projekcích využíval a stále využívá, nicméně ten nerespektuje ani trend, zakomponovaný v časových řadách). Právě na analýze tohoto trendu byla založena tato studie a její výsledky a měla ukázat, že skutečný trend, který se skrývá v empirických datech, poskytuje jiné předpovědi, než používaný deterministický model. Otázkou zůstává, který z přístupů pravděpodobněji poskytne správné budoucí výsledky, které v současnosti neznáme. Bude výhodné pro další pokračování ve výzkumu, pokusit se zkrátit analyzované časové řady na období po druhé světové válce do současnosti a všechny modely přepočítat. Je možné, že výrazné výchylky v období druhé světové války napomohly k tomu, že i v případě modelů ARIMA a náhodné procházky s konstantou došlo k ovlivnění vypočtených předpovědí.

Literatura

- AITKEN, A.C. (1935). On least squares and linear combinations of observations. *Proceedings of the Royal Society of Edinburgh Part A*, vol. 55, pp. 42-47.
- ARLT, J. (1999). „*Moderní metody modelování ekonomických časových řad*“, Grada publishing, 312 s., ISBN 80-7169-539-4.
- ARLT, J., ARLTOVÁ, M. (2007). „*Ekonomické časové řady*“, Grada publishing, 288 s., ISBN 978-80-247-1319-9.
- ARLTOVÁ, M., LANGHAMROVÁ, JI., LANGHAMROVÁ, JA. (2013). Development of life expectancy in the Czech Republic in years 1920-2010 with an outlook to 2050. *Prague Economic Papers*, vol. 22, no. 1, pp. 125-143. ISSN 1210-0455.
- BELL, W. R. (1997). Comparing and assessing time series methods for forecasting age-specific fertility and mortality rates. *Journal of Official Statistics*, 13(3): pp. 279-303.
- BELL, W. R., MONSELL, B. (1991). Using principal components in time series modelling and forecasting of age-specific mortality rates. In: *Proceedings of the American Statistical Association, Social Statistics Section*, pp. 154-159.
- BOOTH, H., TICKLE, L., SMITH, L. (2005). Evaluation of the variants of the Lee-Carter method of forecasting mortality: a multi-country comparison. *New Zealand Population Review*, 31(1), pp. 13-34.
- BOX, G.E.P., JENKINS, G. (1970). „*Time series analysis: Forecasting and control*“, San Francisco, Holden-Day, 537 pp.
- BOX, G.E.P., PIERCE, D.A. (1970). Distribution of the Autocorrelations in Autoregressive Moving Average Time Series Models, *Journal of the American Statistical Association*, vol. 65, pp. 1509-1526.
- COALE, ANSLEY J., KISKER, ELLEN E. (1986). Mortality Crossovers: Reality or Bad Data? *Population Studies*. Vol. 40, pp. 389-401.
- ČSÚ. (2013a). *Úmrtnostní tabulky – metodika*. [on-line] Dostupné z: <http://www.czso.cz/csu/redakce.nsf/i/umrtnostni_tabulky_metodika>
- ČSÚ. (2013b). *Projekce obyvatelstva České republiky do roku 2100*. [on-line] Dostupné z: <<http://www.czso.cz/csu/2013edicniplan.nsf/p/4020-13>>
- ENGLE, R. F., GRANGER, C. W. J. (1987). Co-integration and error correction: Representation, estimation and testing, *Econometrica*, 55(2), pp. 251-276.
- ERBAS, B., ULLAH, S., HYNDMAN, R. J., SCOLLO, M., ABRAMSON, M. (2012). Forecasts of COPD mortality in Australia: 2006-2025. *BMC Medical Research Methodology*, vol. 2012, pp. 12-17.

- FIALA, T. (2006). Dva přístupy modelování vývoje úmrtnosti v populační projekci a jejich aplikace na populaci ČR. In: *Forum Statisticum Slovacum*, 4/2006. Bratislava, 05.10.2006 – 06.10.2006. Bratislava: Slovenská statistická a demografická spoločnosť, 2006, s. 44–55. ISSN 1336-7420.
- GARDNER JR. E.S., MCKENZIE, E. (1985). Forecasting Trends in Time Series. *Management Science*, 31(10), pp. 1237-1246.
- HALLEY, E. (1693). „*An Estimate of the Degrees of the Mortality of Mankind*“.
- HYNDMAN, R. J., KOEHLER, A. B., SNYDER, R. D., GROSE, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18(3), pp. 439–454.
- HYNDMAN, R. J., SHANG, HAN LIN (2009). Forecasting functional time series. *Journal of the Korean Statistical Society*, 38(3), pp. 199-221. (With discussion)
- HYNDMAN, R. J. (2012). *demography: Forecasting mortality, fertility, migration and population data*. R package v. 1.16, Dostupné z: < <http://robjhyndman.com/software/demography/>>
- JARQUE, C.M., BERA, A.K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals, *Economics Letters* 6(3), pp. 255–259.
- KEYFITZ, N. (1991). Experiments in the projection of mortality. *Canadian Studies in Population*, 18(2), pp. 1–17.
- KOSCHIN, F. (2005). How to forecast fertility? In: *Society of Actuaries'05 Annual Meeting & Exhibit*. New York, 14.11.2005 – 16.11.2005. New York: SOA, 2005, s. 1–11.
- KOSCHIN, F. ET AL. (2007). *Prognóza lidského kapitálu obyvatelstva České republiky do roku 2050*. Praha: Oeconomica, 2007. 105 s. (Další autoři: Fiala, T., Fischer, J., Hlavínová, H., Hulík, V., Kačerová, E., Langhamrová, JI., Mazouch, P., Pikálková, S., Šťastnová, P., Fořtlová, S.).
- LEE, R. D., CARTER, L. R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, vol. 87, pp. 659–675.
- LEE R. D., TULJAPURKAR, S. (1994). Stochastic population forecasts for the United States: beyond high, medium, and low. *Journal of the American Statistical Association*, vol. 89, pp. 1175–1189.
- LUNDSTRÖM, H., QVIST, J. (2004). Mortality Forecasting and Trend Shifts: An Application of the Lee-Carter Model to Swedish Mortality Data. *International Statistical Review / Revue Internationale de Statistique*, Vol. 72, No. 1 (Apr., 2004), pp. 37-50.
- MALTHUS, T. R. (1798a). An Essay on the Principle of Population, In: *Oxford World's Classics reprint*, end of Chapter VII., p. 61.
- MALTHUS, T. R. (1798b). An Essay on the Principle of Population. In: *Oxford World's Classics reprint*, Chapter V., pp. 39–45.
- MÉLARD, G., PASTEELS, J.M. (2000). Automatic ARIMA Modeling Including Intervention, Using Time Series Expert Software. *International Journal of Forecasting*, vol. 16, 497-508.
- MURPHY, M. J. (1995). The prospect of mortality: England and Wales and the United States of America, 1962–1989. *British Actuarial Journal*, 1(2), pp. 331–350.
- ORD, K., LOWE, S. (1996). Automatic Forecasting. *The American Statistician*, 50(1), pp. 88-94.
- JEL Classification:** C22, C53, J11

Summary

Modelling of demographic time series for the purposes of the demographic projections

This paper deals with the main differences between the deterministic and stochastic approaches for the calculations of demographic projections. In the past, there were mostly used the deterministic forecasts, which were easy to compute, but currently they are not sufficient for quality explanation of the estimated future trend. Because the

stochastic approach is unfortunately relatively computationally expensive and high demands are expected on the quantity and quality of input data, there will be presented the approach of individual modelling of the time series by Box-Jenkins methodology on the selected demographic indicator – the age-specific death rates of males and females, because the estimated future development of mortality of analysed population is one of the key input information for all population projection. For the purposes of the analysis there will be available the data matrices from the Czech Statistical Office from 1920 to 2012. With use of mentioned methodology there will be constructed the individual optimized ARIMA models (with constant or without), for the persons aged 0–100+ years and in the next part the individual models of random walk with drift, for males and females separately. Based on this models there will be constructed the forecasts of logarithms of age-specific death rates for the period from 2013 to 2050. The obtained forecasts will be confronted and according to their provided results and their computational complexity there will be determined the recommendations about their suitability or their further improvement. By finding of a universal model for all age groups in a particular population we would accelerate the computational process and the possibility of more frequent recalculation of the demographic projections.

Key words: age-specific death rates, time series, ARIMA, random walk, projections.

Porovnání nových přístupů ve shlukování nominálních dat

Zdeněk Šulc

zdenek.sulc@vse.cz

Doktorand oboru statistika

Školitel: Doc. Ing. Iva Pecáková, CSc., (pecakova@vse.cz)

Abstrakt: V poslední době roste poptávka po shlukování kategoriálních dat. Tato oblast není zatím dostatečně prozkoumána, především proto, že je obtížné definovat vzdálenosti mezi objekty charakterizovanými kategoriálními proměnnými. Vzdálenosti mezi objekty se v současnosti určují téměř výhradně za pomoci koeficientu prosté shody, který má mnoho známých nedostatků. Jedním z jeho největších nedostatků je neuvažování důležitých charakteristik v souboru, jako je četnost dané kategorie určité proměnné nebo počet kategorií dané proměnné, které by mohly být použity pro účely definování vzdálenosti. V nedávné době bylo představeno velké množství měr, které mají nedostatky koeficientu prosté shody řešit. Vlastnosti těchto měr však nikdy nebyly důkladně zkoumány, a proto se jejich použití většinou omezilo na text, v němž byly představeny.

Cílem tohoto článku je prozkoumat a ohodnotit vybrané míry podobnosti, které byly v nedávné době navrženy pro účely shlukování objektů charakterizovaných nominálními proměnnými. Tyto míry byly ověřovány na několika datových souborech. Získané shluky jsou hodnoceny podle různých indexů zahrnujících normalizovaný Giniho koeficient a normalizovanou entropii, modifikované pseudo F indexy založené na Giniho koeficientu a na entropii. Výsledná hodnocení jsou porovnávána s výsledky dosaženými použitím koeficientu prosté shody a dále pak s výsledky dvoukrokové shlukové analýzy používané v IBM SPSS a analýzy latentních tříd v programu LatentGold. Provedené experimenty naznačují, že některé z nedávno navržených měr podobnosti poskytují výrazně lepší shluky v porovnání s koeficientem prosté shody.

Klíčová slova: míry podobnosti, shluková analýza, kategoriální data.

Úvod

Shluková analýza patří mezi nejvýznamnější statistické metody. Nabízí široké možnosti využití, např. v průzkumové analýze dat nebo při určování tržních segmentů. Obvykle je založena na mírách podobnosti (vzdálenosti) objektů, kdy nejpodobnější (nejbližší) objekty spolu začínají tvořit shluky. Problematika shlukové analýzy je poměrně dobře prozkoumána v případě, že jsou data tvořena spojitými proměnnými. Potom se pro určování podobnosti mezi objekty používá některá ze známých měr vzdálenosti, nejčastěji *euklidovská vzdálenost*. V současné době vzrůstá poptávka po shlukování kategoriálních dat. Tato oblast není zatím dostatečně prozkoumána, především proto, že je obtížné definovat vzdálenosti mezi objekty charakterizovanými kategoriálními proměnnými. Běžně se používají dva přístupy. První spočívá v transformaci kategoriálních proměnných na skupiny binárních proměnných. Poté je provedena shluková analýza, která využívá měr podobnosti pro binární proměnné, jejichž problematika je poměrně dobře prozkoumána. Nevýhodou tohoto přístupu je možná ztráta důležitých vztahů mezi původními proměnnými. Druhým přístupem, kterým se zabývá tento článek, je použití měr podobnosti, které byly pro kategoriální data přímo navrženy. Takových měr bylo představeno velké množství, ale navzdory jejich zajímavým výsledkům nebyla drtivá většina z nich použita mimo text, ve kterém byly představeny.

Jediná klasická míra, které se hojně používá v praktických aplikacích, je *koeficient prosté shody* (*simple matching coefficient*). Další využívanou mírou je *věrohodnostní míra vzdálenosti* (*log-likelihood distance*), která našla uplatnění ve *dvoukrokové shlukové analýze* v softwaru IBM SPSS. Pro účely shlukování lze v současnosti použít i *modely latentních tříd*, jejichž rozvoj nastal zejména v posledních letech. Zjištěné hodnoty *koeficientu prosté shody* a dále výsledky *dvoukrokové shlukové analýzy* a *analýzy latentních tříd* budou v článku porovnávány s hodnotami měr, které byly navrženy v nedávné době. Kritériem porovnání bude kvalita vytvořených shluků, která bude posuzována na základě vnitroshlukové variability. Míry budou porovnávány na několika datových souborech lišících se počtem pozorování, proměnných a kategorií. Všechny použité vzorce budou vycházet z datové matice $\mathbf{X} = [x_{ic}]$, kde $i = 1, \dots, n$ (n představuje celkový počet objektů) a $c = 1, \dots, m$ (m představuje celkový počet proměnných).

Koeficient prosté shody

Koeficient prosté shody (*simple matching coefficient*), také známý jako míra *overlap*, je nejjednodušší mírou podobnosti. Při zjišťování podobnosti mezi objekty $\mathbf{x}_i$ a $\mathbf{x}_j$ nabývá pro c -tou proměnnou hodnoty 1 v případě jejich shody a hodnoty 0, pokud se proměnné odlišují. Podobnost S_c mezi proměnnými x_{ic} a x_{jc} je tak vyjádřena vztahem

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 1 & \text{jestliže } x_{ic} = x_{jc} \\ 0 & \text{jinak} \end{cases} \quad (1)$$

Míra podobnosti mezi dvěma objekty je spočítána jako aritmetický průměr dílčích měr S_c ,

$$S(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{c=1}^m S_c(x_{ic}, x_{jc})}{m} \quad (2)$$

Každou míru podobnosti $S(\mathbf{x}_i, \mathbf{x}_j)$, zkráceně S_{ij} , lze převést na míru vzdálenosti D_{ij} . V případě míry *overlap* lze tento vztah vyjádřit jako

$$D_{ij} = 1 - S_{ij} \quad (3)$$

Míra *overlap* je široce rozšířena zejména díky své jednoduchosti. Ovšem s jednoduchostí je také spojena jedna z jejích největších nevýhod. Tou je nerozlišování četných a vzácných kategorií u jednotlivých proměnných. Četné i vzácné kategorie mají potom shodnou váhu, čímž se ztrácí část informace, která by mohla pomoci určení, které objekty jsou si podobnější a které vzdálenější.

Nedávno představené míry podobnosti

Míra *Eskin* byla navržena v Eskin et al. (2002) tak, aby přiřazovala vyšší váhy neshodám, které se vyskytují u proměnných s vyšším počtem kategorií. Minimální hodnoty S_c v případě neshody kategorií je potom dosaženo, vyskytují-li se u c -té proměnné pouze dvě kategorie, a maximální hodnoty, vyskytují-li se všechny kategorie této proměnné pouze jednou. Podobnost dvou objektů pro c -tou proměnnou je daná vztahem

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 1 & \text{jestliže } x_{ic} = x_{jc} \\ \frac{n_c^2}{n_c^2 + 2} & \text{jinak} \end{cases} \quad (4)$$

kde n_c počet kategorií u c -té proměnné. Míra podobnosti mezi dvěma objekty pak může být vypočítána podle vztahu (2). Míra vzdálenosti je vyjádřena jako

$$D_{ij} = \frac{1}{S_{ij}} - 1 \quad (5)$$

Míra OF (*occurrence frequency*) je založena na systému vah, který přiřazuje vyšší váhy četnějším kategoriím v případě neshod. Více informací např. v Boriah et al., (2008). Podobnost dvou kategorií je charakterizována vztahem

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 1 & \text{jestliže } x_{ic} = x_{jc} \\ \frac{1}{1 + \ln \frac{n}{f(x_{ic})} \cdot \ln \frac{n}{f(x_{jc})}} & \text{jinak} \end{cases} \quad (6)$$

kde $f(x_{ic})$ vyjadřuje četnost hodnoty x_{ic} u c -té proměnné. S_c dosahuje minimální hodnoty, pokud se x_{ic} i x_{jc} vyskytují právě jednou, maximální hodnoty v případě, že obě hodnoty vyskytují $n/2$ krát. Míra podobnosti lze vypočítat podle vztahu (2) a míra vzdálenosti podle (5).

Míra IOF (*inverse occurence frequency*) má převrácený systém vah vůči míře OF. Neshodám u méně častých kategorií jsou přiřazovány vyšší váhy, naopak, častým kategoriím váhy nižší. Tento postup je vyjádřen vztahem

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 1 & \text{jestliže } x_{ic} = x_{jc} \\ \frac{1}{1 + \ln f(x_{ic}) \cdot \ln f(x_{jc})} & \text{jinak} \end{cases} \quad (7)$$

Míra podobnosti a míra vzdálenosti může být vypočítána podle vztahů (2) respektive (5).

Míra *Lin* byla představena v Lin, (1998). Tato míra přiřazuje vyšší váhy četnějším kategoriím v případě jejich shody a nižší váhy méně častým kategoriím v případě neshod, tedy

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 2 \cdot \ln p(x_{ic}) & \text{jestliže } x_{ic} = x_{jc} \\ 2 \cdot \ln(p(x_{ic}) + p(x_{jc})) & \text{jinak} \end{cases} \quad (8)$$

kde $p(x_{ic})$ vyjadřuje relativní četnost hodnoty x_{ic} u c -té proměnné. Minimální hodnoty v případě shody je dosaženo, pokud se hodnota x_{ic} vyskytuje pouze jedenkrát, a maximální hodnoty, vyskytuje-li se hodnota x_{ic} n krát. Při neshodě je minimální hodnoty dosaženo, pokud se x_{ic} i x_{jc} vyskytují pouze jedenkrát, a maximální hodnoty, pokud se obě hodnoty vyskytují právě $n/2$ krát. Míra podobnosti je vypočítána jako

$$S(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{c=1}^m S_c(x_{ic}, x_{jc})}{\sum_{c=1}^m (\ln p(x_{ic}) + \ln p(x_{jc}))} \quad (9)$$

a míra vzdálenosti podle vztahu (5).

Další možnosti shlukování

Alternativní způsob shlukování dat s kategoriálními proměnnými vychází z *věrohodnostní míry vzdálenosti* (*log-likelihood distance*). Na rozdíl od výše představených měř se tato míra nepoužívá v klasické hierarchické shlukové analýze. Je možné se s ní setkat ve *dvoukrokové shlukové analýze*, která je používána v softwaru IBM SPSS. Tato metoda vychází z upraveného algoritmu BIRCH, který umožňuje použití u spojitých i u kategoriálních proměnných. Více například v (Žambochová, 2010). U spojitých proměnných předpokládá normální rozdělení a u kategoriálních proměnných rozdělení multinomické. V prvním kroku dochází k vytvoření pomocných shluků, které probíhá sekvenčně. S tím je spojena jedna z největších nevýhod této metody, kterou je závislost výsledku shlukování na počátečním pořadí objektů. Ve druhém kroku dochází k aplikaci shlukovacích algoritmů na pomocné shluky. Podrobně popsaný algoritmus *dvoukrokové shlukové analýzy* je možné nalézt např. v SPSS Help (SPSS, Inc. 2013).

Analýza latentních tříd (LCA) představuje modelový přístup shlukování kategoriálních dat. Předpokládá, že data jsou složena z několika homogenních skupin, které jsou v datech promíchány. Vychází proto z předpokladu existence latentní (skryté) kategoriální proměnné, která je složena z diskrétních a vzájemně disjunktních latentních tříd. Za použití pozorovatelných kategoriálních (také manifestních) proměnných je pro každý objekt určena pravděpodobnost příslušnosti do každé latentní třídy. Toho se dá využít při shlukování objektů, kdy je každý objekt přiřazen do shluku (tedy latentní třídy) s nejvyšší pravděpodobností přiřazení. Více informací o LCA ve shlukování lze nalézt např. v (Mičkal, 2007).

Dvoukroková shluková analýza i *analýza latentních tříd* jsou komplexní metody, které umožňují úpravu velkého množství vstupních parametrů (především LCA), proto se jimi vytvořené shluky mohou za použití různého nastavení odlišovat. Z tohoto důvodu bylo u obou metod použito standardní nastavení všech parametrů.

Kritéria hodnocení výsledných shluků

Aplikace uvedených nových měr při shlukování budou vyhodnoceny pomocí měr založených na měření vnitroskupinové variability a měr inspirovaných *pseudo F indexy*, které jsou například vysvětleny v Řezanková et al., (2011).

Vnitroskupinová variabilita je důležitým ukazatelem kvality shlukového řešení. Se vzrůstajícím počtem shluků dochází k poklesu této variability. Za nejlepší míru podobnosti je považována taková, která kvantifikuje nejrychlejší pokles vnitroskupinové variability, protože shluky takové míry jsou nejhomogennější. Pro účely měření vnitroskupinové variability byly vybrány dvě míry, *normalizovaný Giniho koeficient* a *normalizovaná entropie*.

Giniho koeficient, také známý jako *nomvar*, měří variabilitu nominálních proměnných. Lze jej vyjádřit jako

$$G_{gc} = 1 - \sum_{u=1}^{K_c} \left(\frac{n_{gcu}}{n_g} \right)^2, \quad (10)$$

kde G_{gc} je *Giniho koeficient* pro g -tý shluk ($g = 1, 2, \dots, k$; k je počet shluků), n_{gcu} je počet objektů ve shluku g , které u proměnné c ($c = 1, 2, \dots, m$) nabývají kategorie u ($u = 1, 2, \dots, K_c$; K_c je počet kategorií c -té proměnné). Protože *nomvar* standardně nenabývá hodnot od nuly do jedné, je potřeba jej znormalizovat podle vztahu

$$G'(k) = \sum_{g=1}^k \frac{n_g}{n \cdot m} \sum_{c=1}^m \frac{K_c}{K_c - 1} G_{gc}, \quad (11)$$

kde n je celkový počet objektů a m celkový počet proměnných. Čím více se hodnota koeficientu blíží nule, tím je vnitroskupinová variabilita nižší.

Alternativním způsobem měření vnitroskupinové variability nominálních proměnných je *entropie*, která dosahuje nejvyšší hodnoty, jsou-li všechny kategorie shodně zastoupeny a nulové hodnoty, jsou-li všechny objekty právě v jedné kategorii. Lze ji vyjádřit jako

$$H_{gc} = - \sum_{u=1}^{K_c} \left(\frac{n_{gcu}}{n_g} \ln \frac{n_{gcu}}{n_g} \right). \quad (12)$$

Jelikož ani entropie nenabývá hodnot v intervalu od nuly do jedné, je nutné ji znormalizovat podle vztahu

$$H'(k) = \sum_{g=1}^k \frac{n_g}{n \cdot m} \sum_{c=1}^m \frac{H_{gc}}{\ln K_c}. \quad (13)$$

Pseudo F indexy slouží k určení nejlepšího počtu shluků. Na základě počtu objektů, určitého počtu shluků a míry vnitroskupinové variability je určena hodnota statistiky F . Shlukové řešení s nejvyšší hodnotou této statistiky je považováno za nejlepší.

Pro shlukové řešení s k shluky je *pseudo F index založený na Giniho koeficientu* vyjádřen jako

$$I_{\text{PSF-G}}(k) = \frac{(n-k)(G(1) - G(k))}{(k-1)G(k)}, \quad (14)$$

kde $G(1)$ vyjadřuje hodnotu *Giniho koeficientu* v celém datovém souboru a $G(k)$ hodnotu *Giniho koeficientu* při řešení s k shluky.

Pseudo F index založený na entropii lze vyjádřit jako

$$I_{\text{PSF-H}}(k) = \frac{(n-k)(H(1) - H(k))}{(k-1)H(k)}, \quad (15)$$

kde $H(1)$ vyjadřuje entropii v celém souboru a $H(k)$ entropii v řešení s k shluky.

Experimenty

Pro účely porovnávání výše představených měř byly vybrány čtyři datové soubory, které se liší počtem objektů, proměnných a kategorií. Všechny soubory pocházejí z UCI data depository (Bache a Lichman, 2013) a jsou volně dostupné. Tabulka 1 zobrazuje charakteristiky použitých souborů v podobě, jaké vstupovaly do experimentu.

Tabulka 6: Charakteristiky použitých souborů

	Car	Hayes Roth	Breast Cancer	Post Operative
počet objektů	1728	132	683	90
počet proměnných	6	5	9	7
min. počet kat.	3	3	9	2
max. počet kat.	4	4	10	3

Pro výpočet matic vzdáleností podle jednotlivých měř vzdálenosti byl použit *Matlab*. Na základě matic vzdáleností byly v IBM SPSS spočteny hierarchické shlukové analýzy pro dva až šest shluků. Nakonec byly v softwaru *Matlab* spočteny míry hodnotící kvalitu výsledných shluků.

Tabulka 7: Hodnotící kritéria pro soubor Car

	# shluků	1	2	3	4	5	6
OVERLAP	Gnorm	1,000	0,940	0,900	0,878	0,845	0,805
	Hnorm	1,000	0,936	0,899	0,875	0,843	0,803
	PSF-G		102,433	89,205	74,916	73,741	77,581
	PSF-H		103,803	84,709	71,917	70,452	73,320
ESKIN	Gnorm	1,000	0,917	0,833	0,806	0,794	0,780
	Hnorm	1,000	0,903	0,833	0,803	0,792	0,778
	PSF-G		146,894	160,465	128,062	103,850	90,026
	PSF-H		161,131	149,078	121,176	97,295	84,186
OF	Gnorm	1,000	0,986	0,939	0,895	0,881	0,842
	Hnorm	1,000	0,984	0,931	0,887	0,868	0,838
	PSF-G		22,934	57,826	71,130	61,230	68,303
	PSF-H		27,522	69,189	80,430	72,128	73,667
IOF	Gnorm	1,000	0,957	0,931	0,911	0,899	0,840
	Hnorm	1,000	0,946	0,918	0,902	0,889	0,824
	PSF-G		74,191	60,352	52,782	45,795	61,253
	PSF-H		88,665	68,685	55,413	47,749	64,648
LIN	Gnorm	1,000	0,958	0,908	0,866	0,818	0,811
	Hnorm	1,000	0,947	0,899	0,861	0,806	0,796
	PSF-G		72,781	81,909	83,463	89,609	75,249
	PSF-H		87,026	86,443	82,043	90,822	77,725
2STEP	Gnorm	1,000	0,917	0,898	0,867	0,844	0,828
	Hnorm	1,000	0,903	0,879	0,841	0,815	0,797
	PSF-G		146,894	94,657	85,428	76,765	69,603
	PSF-H		161,131	112,553	101,798	92,667	83,815
LCA	Gnorm	1,000	0,945	0,945	0,943	0,921	0,911
	Hnorm	1,000	0,948	0,945	0,944	0,911	0,908
	PSF-G		94,280	49,750	34,574	47,425	35,210
	PSF-H		84,675	49,941	34,352	52,065	37,897

Tabulka 2 shrnuje sledované míry pro soubor *Car*. *Normalizovaný Giniho koeficient* (Gnorm) i *normalizovaná entropie* (Hnorm) klesají u všech měř podobnosti. K nejrychlejšímu poklesu u všech shlukových řešení dochází u míry *Eskin*. O druhé a třetí místo se dělí míry *overlap* a *Lin*. Zatímco *overlap* je

lepší v řešeních s dvěma a třemi shluky, *Lin* prokazuje vyšší pokles vnitroskupinové variability u řešení s větším počtem shluků. Nejhorší výsledky v tomto souboru byly získány za pomoci měr OF a IOF.

Shluky vytvořené za pomoci *dvoukrokové shlukové analýzy* (2STEP) byly poměrně kvalitní, celkově se umístily hned za nejlepší mírou *Eskin*. *Analýza latentních tříd* (LCA) zde vykazuje velmi špatné výsledky. Soubor *Car* byl uměle vytvořen s cílem dosahovat maximální vnitroskupinové variability. LCA si s takovým souborem nedokáže poradit a u všech objektů přiřazuje téměř stejné pravděpodobnosti zařazení do všech shluků. U dvou shluků tedy 0,5, u třech 0,33, atd. Výsledné shluky jsou potom velmi nekvalitní.

Pseudo F indexy slouží k určení optimálního shlukového řešení. Pro takové řešení nabývají míry PSF-G a PSF-H nejvyšších hodnot. Ty jsou v tabulkách označeny tučně. Při pohledu na tabulku 2 je zřejmé, že se optimální řešení u jednotlivých měr poměrně odlišuje. Zatímco *overlap*, IOF, 2STEP a LCA preferují dvoushlukové řešení, míra OF řešení se čtyřmi shluky a míra *Lin* řešení s pěti shluky. U míry *Eskin* nebylo jednoznačně určeno optimální řešení. Zatímco míra PSF-G upřednostňuje řešení se třemi shluky, míra PSF-H volí jako optimální dvoushlukové řešení.

Tabulka 8: Hodnotící kritéria pro soubor Hayes Roth

	# shluků	1	2	3	4	5	6
OVERLAP	Gnorm	0,940	0,868	0,763	0,712	0,609	0,590
	Hnorm	0,931	0,861	0,722	0,672	0,585	0,565
	PSF-G		11,103	14,779	13,701	16,937	14,748
	PSF-H		11,027	18,337	16,492	18,288	16,041
ESKIN	Gnorm	0,940	0,809	0,693	0,668	0,626	0,571
	Hnorm	0,931	0,778	0,657	0,628	0,586	0,531
	PSF-G		19,963	21,724	16,362	15,141	15,745
	PSF-H		23,112	24,530	18,613	17,127	17,894
OF	Gnorm	0,940	0,829	0,772	0,720	0,697	0,674
	Hnorm	0,931	0,768	0,708	0,655	0,631	0,605
	PSF-G		16,830	13,951	13,137	11,233	10,156
	PSF-H		26,467	20,258	18,344	15,562	14,162
IOF	Gnorm	0,940	0,809	0,693	0,637	0,580	0,523
	Hnorm	0,931	0,778	0,657	0,601	0,546	0,495
	PSF-G		19,963	21,724	19,594	19,252	19,913
	PSF-H		23,112	24,530	22,061	21,713	22,015
LIN	Gnorm	0,940	0,830	0,766	0,711	0,643	0,571
	Hnorm	0,931	0,813	0,734	0,684	0,617	0,542
	PSF-G		17,934	15,120	13,906	14,926	16,509
	PSF-H		20,164	18,329	15,934	16,843	18,728
2STEP	Gnorm	0,940	0,829	0,716	0,638	0,574	0,531
	Hnorm	0,931	0,768	0,668	0,606	0,546	0,507
	PSF-G		16,830	20,285	20,730	20,452	19,788
	PSF-H		26,467	25,821	23,834	22,797	21,779
LCA	Gnorm	0,940	0,829	0,717	0,627	0,573	0,516
	Hnorm	0,931	0,768	0,668	0,598	0,545	0,491
	PSF-G		16,830	20,195	21,898	20,773	20,869
	PSF-H		26,467	25,679	24,888	23,259	23,024

Tabulka 3 představuje hodnotící kritéria pro soubor *Hayes Roth*. Nejrychlejší pokles vnitroskupinové variability nastal u míry IOF, která je následována ostatními mírami podobnosti v tomto pořadí: *Eskin*, *Lin*, *overlap* a OF. Zajímavé je, že do tříshlukového řešení podávají míry IOF a *Eskin* shodné výsledky.

Z alternativních shlukovacích metod podává lepší výsledky LCA. Od řešení se čtyřmi shluky jsou její výsledky lepší než u IOF. *Dvoukroková shluková analýza* se celkově umístila za mírou IOF.

Ze zkoumaných měr, OF a *Lin* upřednostňují řešení se dvěma shluky, *Eskin* a IOF řešení se třemi shluky. Na základě shluků vytvořených mírou *overlap*, 2STEP a LCA nelze pomocí *pseudo F indexů* jednoznačně určit optimální shlukové řešení.

V souboru *Breast Cancer*, jehož hodnotící kritéria jsou představena v tabulce 4, si z klasických měr podobnosti nejlépe vede míra *Lin*. Ta je následována ostatními mírami v následujícím pořadí: IOF, *overlap*, OF a *Eskin*.

Tabulka 9: Hodnotící kritéria pro soubor Breast Cancer

	# shluků	1	2	3	4	5	6
OVERLAP	Gnorm	0,727	0,717	0,697	0,572	0,567	0,562
	Hnorm	0,682	0,669	0,649	0,516	0,508	0,500
	PSF-G		9,236	14,413	61,222	47,619	39,810
	PSF-H		12,611	17,434	73,029	57,845	49,272
ESKIN	Gnorm	0,727	0,719	0,700	0,686	0,676	0,670
	Hnorm	0,682	0,667	0,648	0,634	0,622	0,615
	PSF-G		7,483	13,099	13,297	12,688	11,448
	PSF-H		14,614	17,536	17,129	16,325	14,757
OF	Gnorm	0,727	0,682	0,650	0,645	0,642	0,639
	Hnorm	0,682	0,635	0,599	0,592	0,588	0,584
	PSF-G		44,996	40,415	28,820	22,411	18,685
	PSF-H		49,792	47,074	34,482	26,952	22,610
IOF	Gnorm	0,727	0,653	0,599	0,596	0,585	0,558
	Hnorm	0,682	0,612	0,553	0,548	0,533	0,504
	PSF-G		76,790	72,435	49,780	41,104	41,028
	PSF-H		77,235	79,257	55,151	47,154	47,841
LIN	Gnorm	0,727	0,604	0,573	0,562	0,551	0,544
	Hnorm	0,682	0,556	0,519	0,507	0,493	0,486
	PSF-G		138,401	91,009	66,205	53,915	45,350
	PSF-H		154,027	106,551	78,226	65,069	54,483
2STEP	Gnorm	0,727	0,575	0,555	0,540	0,521	0,512
	Hnorm	0,682	0,510	0,488	0,467	0,452	0,440
	PSF-G		179,625	105,491	78,471	67,044	56,787
	PSF-H		229,186	134,959	104,066	86,235	74,566
LCA	Gnorm	0,727	0,575	0,558	0,544	0,530	0,523
	Hnorm	0,682	0,510	0,484	0,469	0,453	0,441
	PSF-G		179,625	102,543	75,842	63,025	52,688
	PSF-H		229,186	139,324	103,097	85,871	74,060

Celkově podává nejlepší výsledky *dvoukroková shluková analýza*. Ta je těsně následována výsledky LCA. Obě alternativní metody shlukování podávají mnohem lepší výsledky než všechny standardní míry podobnosti používané v hierarchické shlukové analýze, což svědčí pro jejich schopnost zkoumat vztahy mezi objekty v komplikovaných souborech.

Míry podobnosti OF a *Lin* a stejně tak 2STEP a LCA preferují řešení se dvěma shluky. Míra *overlap* upřednostňuje čtyřshlukové řešení. Míry *Eskin* a IOF jednoznačně neurčují optimální řešení. Zatímco u míry *Eskin* je možné se rozhodovat mezi řešeními se třemi nebo čtyřmi shluky, u míry IOF se rozhoduje mezi řešeními se dvěma nebo třemi shluky.

Tabulka 5 obsahuje přehled hodnotících kritérií pro soubor *Post Operative*. Nejlepších výsledků dosahuje míra IOF. Na druhém pořadí se umístila míra *Eskin*, která je následována mírou *Lin*. O poslední místo se dělí míry *overlap* a OF.

Dvoukroková shluková analýza vytvořila velmi kvalitní shluky, které v případě řešení se třemi a pěti shluky poskytuje nejlepší shluky vůbec. U řešení se dvěma a šesti shluky má lepší shluky míra IOF. LCA na tomto souboru poskytuje podprůměrné výsledky.

Optimální počet shluků se u jednotlivých měr výrazně odlišuje. Míra IOF společně s *dvoukrokovou shlukovou analýzou* určily jako optimální dvoushlukové řešení, míra *Eskin* řešení se třemi shluky, míra *overlap* řešení se čtyřmi shluky a míra OF pětishlukové řešení. Podle LCA je optimální řešení se dvěma nebo třemi shluky. U míry *Lin* nebylo optimální řešení jednoznačně určeno.

Tabulka 10: Hodnotící kritéria pro soubor Post Operative

	# shluků	1	2	3	4	5	6
OVERLAP	Gnorm	0,746	0,696	0,662	0,608	0,579	0,550
	Hnorm	0,735	0,687	0,653	0,607	0,579	0,550
	PSF-G		6,000	5,491	6,687	6,322	6,163
	PSF-H		5,759	5,520	6,359	6,014	5,943
ESKIN	Gnorm	0,746	0,681	0,581	0,560	0,529	0,485
	Hnorm	0,735	0,686	0,587	0,564	0,535	0,487
	PSF-G		9,597	12,118	9,404	8,808	8,903
	PSF-H		7,709	10,710	8,568	8,118	8,371
OF	Gnorm	0,746	0,710	0,666	0,610	0,565	0,546
	Hnorm	0,735	0,689	0,630	0,577	0,534	0,511
	PSF-G		4,623	5,440	6,319	6,861	6,221
	PSF-H		6,064	7,598	7,895	8,133	7,568
IOF	Gnorm	0,746	0,647	0,609	0,547	0,502	0,440
	Hnorm	0,735	0,657	0,615	0,554	0,505	0,440
	PSF-G		15,955	11,279	11,042	10,759	11,299
	PSF-H		13,896	10,621	10,267	10,332	10,666
LIN	Gnorm	0,746	0,698	0,611	0,578	0,535	0,502
	Hnorm	0,735	0,697	0,611	0,576	0,527	0,495
	PSF-G		6,849	8,813	7,806	8,032	7,905
	PSF-H		5,648	7,713	7,239	7,965	7,837
2STEP	Gnorm	0,746	0,666	0,597	0,534	0,487	0,461
	Hnorm	0,735	0,648	0,579	0,516	0,473	0,445
	PSF-G		11,047	11,045	10,855	10,700	9,990
	PSF-H		12,479	12,025	11,490	10,856	10,347
LCA	Gnorm	0,746	0,688	0,629	0,596	0,544	0,535
	Hnorm	0,735	0,666	0,613	0,574	0,534	0,504
	PSF-G		7,764	8,067	7,109	7,814	6,696
	PSF-H		9,542	8,805	8,006	7,993	7,858

Tabulka 6 zobrazuje pořadí, kterého v daném souboru dosáhly zkoumané míry podobnosti nebo metody. Poslední sloupec tabulky znázorňuje průměrné umístění dané míry. Průměrně nejlepší výsledky byly dosaženy za pomoci *dvoukrokové shlukové analýzy*. Tato metoda se napříč všemi soubory umístila nejhůře na třetím místě. To z ní činí velmi dobrou volbu pro doposud neprozkoumané soubory. Na celkově druhém místě se umístila míra IOF. Ta podávala velmi dobré výsledky, především v souborech s jednodušší strukturou. Míra *Eskin* vykazuje průměrné výsledky, avšak s poměrně velkou variabilitou, což ji činí značně nepředvídatelnou. Obdobná situace je u *analýzy latentních tříd*. Míra *Lin* dosahovala lepších výsledků na komplikovanějších souborech. Míry *overlap* a *OF* neexcelovaly v žádném z použitých souborů, a proto jsou jejich výsledky hodnoceny jako nejhorší.

Tabulka 11: Srovnání zkoumaných souborů

	Car	Hayes Roth	Breast Cancer	Post Operative	Průměr
OVERLAP	3,5	6	5	6,5	5,3
ESKIN	1	4	7	3	3,8
OF	5,5	7	6	6,5	6,3
IOF	5,5	2	4	1,5	3,3
LIN	3,5	5	3	4	3,9
2STEP	2	3	1	1,5	1,9
LCA	7	1	2	5	3,8

Závěr

V tomto článku byly porovnány čtyři nedávno představené míry podobnosti, které byly primárně navrženy pro využití v hierarchické shlukové analýze. Tyto míry byly na různých datových souborech porovnávány s běžně používanou mírou *overlap* a dále s výsledky alternativních metod shlukování v podobě *dvoukrokové shlukové analýzy* a *analýzy latentních tříd*.

Téměř všechny nedávno představené míry, kromě míry OF, si vedly lépe než standardně používaná míra *overlap*. Ta ve všech prozkoumaných souborech dosahovala podprůměrných výsledků. Především u menších souborů si nejlépe vedla míra IOF, která přiřazuje neshodám u méně častých kategorií vyšší váhy a v případě častých kategorií váhy nižší. Míra *Lin* vykazovala dobré výsledky u souborů s větším počtem proměnných a kategorií.

Dvoukroková shluková analýza, jako jediná ze všech zkoumaných metod, poskytovala velmi dobré shluky ve všech souborech. Ačkoliv v dílčích případech vycházejí výsledky s použitím některé z nedávno představených měr podobnosti lépe, výhodou této metody je právě dobrá kvalita výsledků napříč datovými soubory, která z ní činí univerzální nástroj pro shlukování kategoriálních dat.

Analýza latentních tříd poskytuje v některých souborech velmi dobré výsledky, avšak v jiných souborech zcela selhává. Při volbě této metody je proto vhodné porovnání jejích výsledků buď s výsledky použití některé ze standardních měr podobnosti v hierarchické shlukové analýze nebo lépe s výsledky *dvoukrokové shlukové analýzy*.

V dalším výzkumu plánuji prozkoumat další skupinu měr podobnosti, která je představena v (Morlini a Zani, 2012) a která slibuje výrazně lepší kvalitu výsledných shluků, převážně u kategoriálních dat se závislými proměnnými.

Literatura

BORIAH, S., CHANDOLA, V., KUMAR, V. 2008. *Similarity measures for categorical data: A comparative evaluation*. In Proceedings of the 8th SIAM International Conference on Data Mining. SIAM, 2008, p. 243-254.

BACHE, K. a LICHMAN, M. (2013): UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science.

ESKIN, E., ARNOLD, A., PRERAU, M., PORTNOY, L., STOLFO, S., V. 2002. *A geometric framework for unsupervised anomaly detection*. In D. Barbar'a and S. Jajodia, editors, Applications of Data Mining in Computer Security, p. 78–100.

LIN, D. 1998. *An information-theoretic definition of similarity*. In ICML '98: Proceedings of the 15th International Conference on Machine Learning. San Francisco: Morgan Kaufmann Publishers Inc., 1998, p. 296-304.

MIČKAL, T. 2007. *Využití modelů latentních tříd při segmentaci trhu*. Diploma thesis.

MORLINI, I., ZANI, S. 2012. *A new class of weighted similarity indices using polytomous variables*. In Journal of Classification, 2012, vol. 29, iss. 2, p. 199-226.

SPSS, Inc. 2013. Help. Chicago, IL : SPSS, Inc., 2013.

ŘEZANKOVÁ, H., LÖSTER, T., HÚSEK, D. 2011. *Evaluation of categorical data clustering*. In Advances in Intelligent Web Mastering – 3. Berlin : Springer Verlag, 2011, p. 173-182.

ŽAMBOCHOVÁ, M. 2010. *Shluková analýza rozsáhlých souborů dat: nové postupy založené na metodě k-průměrů*. Dissertation thesis.

JEL Classification: C38

Summary

Comparison of the new Approaches in Nominal Data Clustering

In recent years, there has been a growing demand for categorical data clustering. This area is not explored sufficiently, mainly because it is difficult to define distances between objects characterized by categorical variables. Distances between objects are almost solely determined by the simple matching coefficient, which has many known drawbacks. One of the greatest, it is not considering the important characteristic of the file, such as frequency of given category of a variable or number of categories of a given variable. These characteristics could be used for purposes of defining the distance. It has been introduced many similarity measures with the aim to deal

with drawbacks of the simple matching coefficient. However, their properties have never been examined properly and therefore their usage has been mostly limited to the article, which they have been introduced in.

The contribution evaluates some of recently introduced similarity measures, which have been proposed for purposes of hierarchical cluster analysis with nominal variables. These measures have been tested on various datasets. Obtained clusters are evaluated by several indices, which include the normalized Gini coefficient and the normalized entropy, the modified pseudo F indices based on the Gini coefficient and the entropy. The evaluated similarity measures are compared to results obtained from the simple matching coefficient and further to results from two-step cluster analysis in IBM SPSS and latent class analysis in LatentGold. The results indicate that some of recently proposed similarity measures provide significantly better clusters in a comparison to commonly used measures, especially to the simple matching coefficient.

Key words: similarity measures, cluster analysis, categorical data.